



PI Legrain P, Benarous R, Blot G, Lassot I;  
 XX WPI; 2002-508795/54.  
 XX

PT Protein-protein complexes for screening drugs or agents that modulate  
 interaction of proteins, e.g. for identifying the Selected Interacting  
 PT Domains (SID), comprises interaction between beta-TrCP and Ras SFI -  
 XX  
 XX Example 9; Page -: 84pp; English.

CC The present invention relates to a new complex of protein-protein  
 CC interaction between betaTrCP (not defined in specification) and Ras SFI.  
 CC The protein-protein complex of the invention is useful for screening  
 CC drugs or agents that modulate interaction of proteins. In particular,  
 CC the protein complex is useful for identifying the Selected Interacting  
 CC Domains (SID). The modulating compounds detected can be used for  
 CC treating tumours. The polynucleotides encoding the protein complex may  
 CC be used in gene therapy. The present amino acid sequence represents  
 CC human beta TrCP (bTrCP) mutant protein #1, as described in the methods  
 CC of the invention.  
 CC Note: This sequence is not shown in the specification but is derived from  
 CC the wild-type human bTrCP sequence (AAU98087) shown on page 47 of the  
 CC specification.  
 CC  
 XX  
 XX Sequence 76 AA:

Query Match 100.0%; Score 240; DB 23; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-23;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 ALPARGLDHIAENILSYDPAKSLCAELVCKEMRYVTSIDGMLMK 45  
 DB 4 ALPARGLDHIAENILSYDPAKSLCAELVCKEMRYVTSIDGMLMK 48

RESULT 2

AAU98087

ID AAU98087 standard; Protein: 219 AA.

AC AAU98087;

DT 24-SEP-2002 (first entry)

DE Human beta TrCP (bTrCP) bait protein sequence.

XX Human; beta TrCP; bTrCP; protein-protein interaction complex; Ras SFI;  
 XX drug screening; selected interacting domain; SID; tumour; gene therapy;  
 XX bait protein; cytosolic.

OS Homo sapiens.

PN WO200250261-A2.

XX 27-JUN-2002.

PF 18-DEC-2001; 2001MO-EP15414.

PR 18-DEC-2000; 2000US-256276P.

XX (HYBR-) HYBRIGENICS.

PA (INRM) INSR NAT SANTE & RECH MEDICALE.

PI Legrain P, Benarous R, Blot G, Lassot I;

XX WPI; 2002-508795/54.

DR N-PSDB; ABR6903.

PT Protein-protein complexes for screening drugs or agents that modulate  
 interaction of proteins, e.g. for identifying the Selected Interacting  
 PT Domains (SID), comprises interaction between beta-TrCP and Ras SFI -  
 XX  
 XX Claim 1; Page 47; 84pp; English.

CC The present invention relates to a new complex of protein-protein  
 CC interaction between betaTrCP (not defined in specification) and Ras SFI.  
 CC The protein-protein complex of the invention is useful for screening  
 CC drugs or agents that modulate interaction of proteins. In particular,  
 CC the protein complex is useful for identifying the Selected Interacting  
 CC Domains (SID). The modulating compounds detected can be used for  
 CC treating tumours. The polynucleotides encoding the protein complex may  
 CC be used in gene therapy. The present amino acid sequence represents the  
 CC human beta TrCP (bTrCP) protein that was used in the methods of the  
 CC invention as a bait protein.  
 CC  
 XX  
 XX Sequence 219 AA:

Query Match 100.0%; Score 240; DB 23; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 3.0e-23;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 ALPARGLDHIAENILSYDPAKSLCAELVCKEMRYVTSIDGMLMK 45  
 DB 147 ALPARGLDHIAENILSYDPAKSLCAELVCKEMRYVTSIDGMLMK 191

RESULT 3

AA848289

ID AA848289 standard; Protein: 448 AA.

AC AA848289;

DT 02-APR-2001 (first entry)

DE Beta-TrCP-N/SKP2.C protein.

XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;  
 XX CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; C-MYC; MDM2; p53; Bax;  
 XX Bad; Bcl-2; tumour; cytosolic.

OS Synthetic.

PN WO200075184-A1.

XX 14-DEC-2000.

PF 05-JUN-2000; 2000MO-US15449.

PR 04-JUN-1999; 99US-0137494.

PA (UYVA) UNIV YALE.

PI Zhang H, Tsvetkov LM, Kondo T;

XX WPI; 2001-061703/07.

DR N-PSDB; AAC84601.

PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
 PT involves altering levels of proteins such as S-phase kinase associated  
 PT proteins 1, 2 and cullin/CDC53 proteins -  
 XX  
 XX Claim 5; Page 109-111; 162pp; English.

CC The invention relates to methods of altering the polypeptide levels in a  
 CC cell, using proteins selected from S-phase kinase associated proteins 1  
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
 CC cullin/CDC53 family of proteins). The method is useful for altering the  
 CC level of p27, cyclin E, Max, Mad, C-MYC, MDM2, p53, Bax, Bad or Bcl-2  
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for  
 CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
 CC that modulate interactions between SKP and target proteins are useful for  
 CC treating tumours.

XX Sequence 448 AA;

Query Match 100.0%; Score 240; DB 22; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-23;



Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPARGIDHIAENILSYLDKSLCAELVCKEWMYRTSGMLMK 45  
 |||  
 Db 121 ALPARGIDHIAENILSYLDKSLCAELVCKEWMYRTSGMLMK 165

RESULT 4  
 AAY24054  
 ID AAY24054 standard; Protein: 569 AA.

AC AAY24054;

XX 30-SEP-1999 (first entry)

DE A human beta-transducin repeat containing protein.

KM Beta-transducin repeat containing protein; beta-Trcp; Skrip;  
 KM proteosome degradation pathway; Vpu protein; beta-catenin; Ikappab;  
 KM human immune deficiency virus-1; HIV-1; cellular protein; Alzheimer's;  
 KM ubiquitinylation; phosphorylated protein; tumour; apoptosis; Alzheimer's;  
 KM antiviral; antitumour; cell cycle regulation; protein degradation;  
 KM and anti-inflammatory; osteo-articular inflammation; acute inflammation;  
 KM tumour necrosis factor.

OS Homo sapiens.

XX Location/Qualifiers

FT Key 147..191

FT Region /note="F box sequence"

FT Region 259..292

FT Region /note="WD motif"

FT Region 304..332

FT Region /note="WD motif"

FT Region 343..372

FT Region /note="WD motif"

FT Region 387..415

FT Region /note="WD motif"

FT Region 427..455

FT Region /note="WD motif"

FT Region 467..492

FT Region /note="WD motif"

FT Region 516..544

FT Region /note="WD motif"

XX WO938969-A1.

XX 05-AUG-1999.

XX 29-JAN-1999: 99WO-FR00196.

XX 09-DEC-1998: 98FR-0015545.

XX 30-JAN-1998: 98FR-0001100.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX (INSP ) INST PASTEUR.

XX Arenzana Seisdedos F, Benarous R, Concorde J, Durand H;

XX Kroll M, Margottin F;

XX WPI: 1999-469329/39.

XX N-PSDB: AAX86501.

XX New human beta-transducin repeat containing protein and its

XX fragments useful as, or to screen for, antiviral, antitumour,

XX anti-inflammatory and anti-Alzheimer's agents

XX Claim 1: Page 60-61; 71pp; French.

XX The present sequence represents a human beta-transducin repeat containing

XX protein (beta-Trcp). The protein directs proteins to the proteosome

XX degradation pathways. The protein is able to interact with the Vpu

XX protein of human immune deficiency virus-1 (HIV-1), cellular proteins

CC Ikappab or beta-catenin (bc) and/or protein Skrip. The protein controls  
 CC ubiquitinylation of phosphorylated proteins and thus their targeting to  
 CC proteosomes for degradation. Depending on whether the process is  
 CC inhibited or promoted, the result may be delayed breakdown of CD4 (in  
 CC cases of HIV-1 infection); increased degradation of mutant bc in tumour  
 CC activity of NFkappab; increased degradation of mutant bc in tumour  
 CC cells, or increased bc survival (and reduced apoptosis) in Alzheimer's  
 CC patients. The beta-Trcp protein, and its active peptide fragments, or its  
 CC nucleic acid, are used to screen for anti HIV-1 agents (antivirals),  
 CC antitumour agents that disrupt cell cycle regulation or protein  
 CC degradation in human tumour cells, and anti-inflammatory agents that  
 CC disrupt activation by NFkappab. Fragments of the protein are also  
 CC useful for treating osteo-articular inflammation or acute inflammation  
 CC associated with release of tumour necrosis factor.

XX Sequence 569 AA;

XX Query Match 100.0%; Score 240; DB 20: Length 569;

XX Best Local Similarity 100.0%; Pred. No. 1.1e-22;

XX Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPARGIDHIAENILSYLDKSLCAELVCKEWMYRTSGMLMK 45  
 |||  
 Db 147 ALPARGIDHIAENILSYLDKSLCAELVCKEWMYRTSGMLMK 191

RESULT 5

AB12812  
 ID AAB12812 standard; Protein: 569 AA.

AC AAB12812;

XX 27-NOV-2000 (first entry)

DE Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.

XX Ubiquitin ligase SCF complex; F-box protein; ubiquitinylation; Ikappab;

XX beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;

XX gene therapy; colon cancer; beta-transducin repeat containing protein;

XX beta-Trcp.

XX Mus musculus.

XX JP2000166542-A.

XX 20-JUN-2000.

XX 02-DEC-1998: 98JP-0343437.

XX 02-DEC-1998: 98JP-0343437.

XX (KAGA-) KAGAKU GIYUTSU SHINKO JIGYODAN.

XX WPI: 2000-485550/43.

XX N-PSDB: AAV3131.

XX F-box protein of ubiquitin ligase SCF complex which promotes the

XX ubiquitinylation of Ikappab or beta-catenin

XX Claim 2: Page 9-10; 19pp; Japanese.

XX The present invention describes an F-box motif protein of ubiquitin

XX ligase SCF complex which promotes the ubiquitinylation of Ikappab or

XX beta-catenin and is constituted by Skp1 protein, Cull1 protein and a

XX complex (SCF complex) of F-box protein containing F-box motif and WD40

XX repeat motif and has the amino acid sequence of 45 residues (AAB12811)

XX or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin

XX ligase FWD1 protein) and (AAB12813, which is human beta-transducin

XX repeat containing protein (beta-Trcp)). The F-box protein can be used for

XX the gene therapy of colon cancer by being recombinant to a virus vector.

XX Sequence 569 AA;

Query Match 100.0%; Score 240; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIADENILSYLDKSLCAAEIVCKEYRVTSQGLMKK 45  
 DB 147 ALPARGLDHIADENILSYLDKSLCAAEIVCKEYRVTSQGLMKK 191

RESULT 6  
 AAB12813

ID AAB12813 standard; protein; 569 AA.

AC AAB12813;

DT 27-NOV-2000 (first entry)

DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.

KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;  
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KW beta-TrCP.

OS Homo sapiens.

PN JP2000166542-A.

PD 20-JUN-2000.

PF 02-DEC-1998; 98JP-0343437.

PR 02-DEC-1998; 98JP-0343437.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

DR WPI: 2000-485550/43.

DR N-PSDB; AAAV3132.

PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of Ikappab or beta-catenin

PS Claim 3; Page 10-12; 19pp; Japanese.

CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of Ikappab or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12813)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.

SQ Sequence 569 AA;

Query Match 100.0%; Score 240; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIADENILSYLDKSLCAAEIVCKEYRVTSQGLMKK 45  
 DB 147 ALPARGLDHIADENILSYLDKSLCAAEIVCKEYRVTSQGLMKK 191

RESULT 7

ID AAY96697 standard; protein; 569 AA.

AC AAY96697;

DT 26-SEP-2000 (first entry)

DE Human beta-TrCP.

KW E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KW anti-inflammatory; immunosuppressive; cytostatic.

OS Homo sapiens.

PN WO200034447-A2.

PD 15-JUN-2000.

PF 10-DEC-1999; 99WO-US29371.

PR 10-DEC-1998; 98US-0210060.

PA (SIGN-) SIGNAL PHARM INC.  
 (VISS) VISSUM RES & DEV CO.

PI Manning AM, Mercurio F, Amit S, Ben-zur Y, Davis M, Hatzubai A,  
 PI Layon I, Yaron A;

DR WPI: 2000-431294/37.  
 DR N-PSDB; AAA51229.

PT Polypeptide enhancing phosphorylated Ikappab ubiquitination useful for  
 PT treating disorder associated with NF-kappa-B activation e.g. cancer,  
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
 PT variant

PS Claim 21; Page 72-74; 77pp; English.

CC Human beta-TrCP, an F-box/WD protein family member, has been shown to  
 CC have homology to human E3 ubiquitin ligase (E3). E3 enhances  
 CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of  
 CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B  
 CC degradation via the ubiquitin pathway is useful for identifying  
 CC modulators of this process for use in treating diseases associated with  
 CC activation of NF-kappa-B. In vitro analysis suggests that deletion of  
 CC the F-box results in a protein that functions as a dominant negative  
 CC deletion mutant. In a protein that functions as a dominant negative  
 CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated  
 CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B  
 CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat  
 CC inflammatory diseases, autoimmune diseases, cancer and viral infections.

SQ Sequence 569 AA;

Query Match 100.0%; Score 240; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIADENILSYLDKSLCAAEIVCKEYRVTSQGLMKK 45  
 DB 147 ALPARGLDHIADENILSYLDKSLCAAEIVCKEYRVTSQGLMKK 191

RESULT 8

ID AAY83041 standard; protein; 569 AA.

AC AAY83041;

DT 16-AUG-2000 (first entry)

DE F-box protein FBP-1.

KW F-box protein: FBP; diagnosis; treatment; screening; agonist;  
 KW antagonist; proliferative disorder; differentiative disorder;  
 KW breast cancer; prostate cancer; ovarian cancer; cancer;  
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;  
 KW inflammatory disorder; human.

XX OS Homo sapiens.  
 XX PA WO200012679-A1.  
 XX PN 09-MAR-2000.  
 XX PD 27-AUG-1999; 99WO-US19560.  
 XX PF 28-AUG-1998; 98US-0098355.  
 XX PR 03-FEB-1999; 99US-0118568.  
 XX PR 15-MAR-1999; 99US-0124449.  
 XX PA (UNIV ) UNIV NEW YORK STATE.  
 XX PI Chlaur DS, Pagano M, Latres E;  
 DR WPI: 2000-256635/22.  
 DR N-PSDB; AA293350.  
 PS Novel nucleic acid for screening compounds useful for treating  
 PT proliferative and differentiative disorders such as cancer and immune  
 PT disorders comprises sequences encoding ubiquitin ligases  
 CC Disclosure: Figure 3a; 245pp; English.  
 XX Nucleic acids encoding substrate-targeting subunits of ubiquitin  
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis  
 CC of proliferative and differentiated related disorders by measuring  
 CC FBP gene expression. Cells expressing such proteins or  
 CC their fragments are useful for screening compounds. The compounds  
 CC are agonists or antagonists, which are useful for treating a,  
 CC proliferative or differentiative disorder in a mammal, such as  
 CC breast, ovarian and prostate cancer and small cell lung carcinoma  
 CC and also major opportunistic infections, immune disorders,  
 CC cardiovascular diseases and inflammatory disorders. FBP protein,  
 CC analogs, derivatives and their subsequences, anti-FBP antibodies  
 CC are also useful in diagnosis of the disorders.  
 XX SO Sequence 569 AA:  
 Query Match 100.0%; Score 240; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1,1e-22;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 ALPARGLDHIAENILSYIDAKSLCAELVCKEWRVYTSQGLMKK 45  
 DB 147 ALPARGLDHIAENILSYIDAKSLCAELVCKEWRVYTSQGLMKK 191  
 RESULT 9  
 AAY83250  
 ID AAY83250 standard; Protein: 569 AA.  
 XX AAY83250;  
 AC 16-AUG-2000 (first entry)  
 DT 16-AUG-2000 (first entry)  
 XX F-box protein hBetarCP.  
 DE Ubiquitin ligase; SCF; F-box protein; targeted degradation;  
 XX destabilisation; proteolysis; drug discovery; gene therapy; cancer;  
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;  
 KW human.  
 XX Homo sapiens.  
 XX OS  
 XX PA WO200022110-A2.  
 XX PN 20-APR-2000.  
 XX PD 08-OCT-1999; 99WO-US23705.  
 XX PF

PR 09-OCT-1998; 98US-0103787.  
 XX (HARD ) HARVARD COLLEGE.  
 XX PI Zhou P, Howley P;  
 XX WPI: 2000-317970/27.  
 DR N-PSDB; AA293710.  
 DR Targeting degradation of polypeptide useful for treating cancer and  
 PT other proliferative disorders, involves conjugating polypeptide with  
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic  
 PT compound  
 XX Claim 9; page 171; 185pp; English.  
 PS The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin  
 CC ligases) which can be used for the targeted degradation of a target  
 CC polypeptide in vivo. Targeted degradation is achieved by expressing  
 CC the ubiquitin ligase in a cell linked to the interaction domain of  
 CC the target polypeptide and thereby recruiting the target polypeptide  
 CC to the ubiquitin ligase. Such methods are useful for decreasing or  
 CC increasing the level of a target polypeptide and for creating and  
 CC expressing a destabilized polypeptide which is subjected to SCF  
 CC mediated proteolysis. Degrading any desired protein in a cell is  
 CC useful for preventing or treating diseases caused by the presence of  
 CC abnormal amount of the specific polypeptides, for drug discovery and  
 CC for gene therapy. Diseases treated include cancer, by degradation of  
 CC oncoproteins, Huntington's disease, other proliferative disorders and  
 CC microbial infections. The method provides a quick and easy  
 CC alternative to gene knockout technology. The target polypeptide can  
 CC be degraded at all stages, or a specific stage, of development in the  
 CC mature animal.  
 XX SO Sequence 569 AA:  
 Query Match 100.0%; Score 240; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1,1e-22;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 ALPARGLDHIAENILSYIDAKSLCAELVCKEWRVYTSQGLMKK 45  
 DB 147 ALPARGLDHIAENILSYIDAKSLCAELVCKEWRVYTSQGLMKK 191  
 RESULT 10  
 AAY83254  
 ID AAY83254 standard; Protein: 569 AA.  
 XX AAY83254;  
 AC 16-AUG-2000 (first entry)  
 DT 16-AUG-2000 (first entry)  
 XX F-box protein FWDLP.  
 DE Ubiquitin ligase; SCF; F-box protein; targeted degradation;  
 XX destabilisation; proteolysis; drug discovery; gene therapy; cancer;  
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;  
 KW mouse; ss.  
 XX Mus musculus.  
 XX OS  
 XX PA WO200022110-A2.  
 XX PN 20-APR-2000.  
 XX PD 08-OCT-1999; 99WO-US23705.  
 XX PF 09-OCT-1998; 98US-0103787.  
 XX PR (HARD ) HARVARD COLLEGE.  
 XX PA Zhou P, Howley P;  
 XX PI

The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin ligases) which can be used for the targeted degradation of a target polypeptide in vivo. Targeted degradation is achieved by expressing the ubiquitin ligase in a cell linked to the interaction domain of the target polypeptide and thereby recruiting the target polypeptide to the ubiquitin ligase. Such methods are useful for decreasing or increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF-mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating any diseases caused by the presence of abnormal amount of the specific polypeptides, for drug discovery and for gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

SQ Sequence 569 AA;

Query Match	100.0%;	Score 240;	DB 21;	Length 569;
Best Local Similarity	100.0%;	Pred. No. 1,1e-22;		
Matches				

[illegible]

Db 147 ALPARGLDHIAENILSYLDAKSLCAELVCKEWMYRVTS DGM L W K K 45

RESULT 11  
AAV44340

AAV44249 standard; protein; 569 AA.

AC YY  
AA44249;

28-FEB-2000 (first entry)

human cell signalling protein-12

inflammatory disorder: cirrhosis, cell proliferation; CSIGP-12; cell

...disease; multiple sclerosis.

Homo sapiens.

Modified-site  
19

Modified-site	39	potential phosphorylation site <sup>a</sup>
ST		
ST		
ST		

Modified-site	91	phosphorylation site"
Modified-site	91	phosphorylation site"

Modified-site	109	phosphorylation site

Modified-site	Free-radical site
162	

modified-site	266	potential site
modified-site	266	potential site

/note= "Potential about 1000 ft. from shore  
288

020 /note="Potential phosphorus"

FT	Modified-site	376	/note= "Potential phosphorylation site"
FT	Modified-site	381	/note= "Potential phosphorylation site"
FT	Modified-site	411	/note= "Potential phosphorylation site"
FT	Modified-site	418	/note= "Potential phosphorylation site"
FT	Modified-site	451	/note= "Potential phosphorylation site"
FT	Modified-site	514	/note= "Potential phosphorylation site"
FT	Modified-site	519	/note= "Potential phosphorylation site"
FT	Modified-site	535	/note= "Potential phosphorylation site"
FT	Modified-site	536	/note= "Potential phosphorylation site"
FT	Modified-site	17	/note= "Potential phosphorylation site"
FT	Modified-site	77	/note= "Potential glycosylation site"
FT	Modified-site	416	/note= "Potential glycosylation site"
FT	Modified-site	320..334	/note= "Potential glycosylation site"
FT	Region	360..374	/label= Signature_sequence
FT	Region	403..417	/label= Signature_sequence
FT	Region	443..457	/label= Signature_sequence
FT	Region	483..497	/label= Signature_sequence
FT	Region	532..546	/label= Signature_sequence
PN	W09958558-A2.		
XX	18-NOV-1999.		
XX	13-MAY-1999;	99MO-US10567.	
PR	13-MAY-1998;	98US-0085343.	
PR	26-AUG-1998;	98US-0098010.	
PA	(INCY-) INCYTE PHARM INC.		
PI	Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;		
PI	Baughn MR, Yang J;		
DR	WPI; 2000-086432/07.		
DR	N-PSDB; AA29233.		
XX	Human cell signaling proteins useful for, e.g. diagnosing cell		
XX	proliferative and inflammatory disorders -		
PS	Claim 1; Page 77-78; 90pp; English.		
CC	The present sequence is cell signalling protein-12 (CSIGP-12) encoded		
CC	by cDNA obtained from incyte clone 3239149 of COLAUC701 library. It is		
CC	expressed in musculoskeletal, gastrointestinal and nervous tissues and is		
CC	found to be homologous to beta-transducin repeats containing		
CC	protein. Fragments of CSIGP encoding nucleic acid can be used as		
CC	hybridisation probe for detecting CSIGP related sequences or allelic		
CC	variants. Recombinant CSIGP can be produced in host cells by transforming		
CC	them with genetically engineered vectors. Agonists or antagonists can be		
CC	used in the treatment of cell proliferative and inflammatory disorders		
CC	associated with decreased or increased CSIGP expression. CSIGP is used in		
CC	the diagnosis, prevention and treatment of cell proliferative disorders		
CC	like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory		
XX	disorders like AIDS, Addison's disease, multiple sclerosis, etc.		

```
SQ Sequence 569 AA;
Query Match 100.0%; Score 240; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 1,1e-22;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPARGLDHIAENILSYLDKSLCAAEIVCKEMRYVTSQGLMKK 45;
    |||
DB 147 ALPARGLDHIAENILSYLDKSLCAAEIVCKEMRYVTSQGLMKK 191

RESULT 12
AAB48298
ID AAB48298 standard; protein; 569 AA.
XX
AC AAB48298;
XX
DT 02-APR-2001 (first entry)
XX
DE Human ZF11 protein.
XX
KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
KW Bad; Bcl-2; tumour; cytostatic.
XX
OS Homo sapiens.
XX
PN M0200075184-A1.
XX
PD 14-DEC-2000.
XX
PF 05-JUN-2000; 2000MO-U0515449.
XX
PR 04-JUN-1999; 990S-0137494.
XX
PI (UYVA) UNIV YALE.
XX
PI Zhang H, Tsvetkov LM, Kondo T;
XX
DR WPI: 2001-061703/07.
XX
DR N-PSDB; AAC84610.
XX
PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
PT involves altering levels of proteins such as S-phase kinase associated
PT proteins 1, 2 and cullin/CDC53 proteins -
XX
PS Claim 3; Page 130-132; 162pp; English.
XX
CC The invention relates to methods of altering the polypeptide levels in a
CC cell, using proteins selected from S-phase kinase associated proteins 1
CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
CC cullin/CDC53 family of proteins). The method is useful for altering the
CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents
CC that modulate interactions between SKP and target proteins are useful for
CC treating tumours.
XX
SQ Sequence 569 AA;
Query Match 100.0%; Score 240; DB 22; Length 569;
Best Local Similarity 100.0%; Pred. No. 1,1e-22;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPARGLDHIAENILSYLDKSLCAAEIVCKEMRYVTSQGLMKK 45
    |||
DB 147 ALPARGLDHIAENILSYLDKSLCAAEIVCKEMRYVTSQGLMKK 191

RESULT 13
AAM78583
ID AAM78583 standard; protein; 579 AA.
XX
AC AAM78583;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1245.
XX
KW Human: cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN M0200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001MO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZM;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
XX
DR N-PSDB; AAK51716.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 3504-3505; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 579 AA;
Query Match 100.0%; Score 240; DB 22; Length 579;
Best Local Similarity 100.0%; Pred. No. 1,1e-22;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPARGLDHIAENILSYLDKSLCAAEIVCKEMRYVTSQGLMKK 45
    |||
DB 157 ALPARGLDHIAENILSYLDKSLCAAEIVCKEMRYVTSQGLMKK 201

RESULT 14
AAM00847
ID AAM00847 standard; protein; 590 AA.
XX
AC AAM00847;
```

```

XX 01-OCT-2001 (first entry)
DT
XX
DE Human bone marrow protein, SEQ ID NO: 210.
XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
OS Homo sapiens.
XX
PN WO200153453-A2.
XX
PD 26-JUL-2001.
XX
PF 23-DEC-2000; 2000MO-US34960.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0663191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX
DR WPI: 2001-488707/53.
DR N-PSDB: AAH89966.
XX
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
PS Claim 10; Page 354-355; 648bp; English.
XX
CC The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, or may result from a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
SQ Sequence 590 AA;

```

```

Query Match 100.0%; Score 240; DB 22; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALPARGDHIAENILSYLDAKSLCAELVCKEWRVTSQGLMKK 45
Db 168 ALPARGDHIAENILSYLDAKSLCAELVCKEWRVTSQGLMKK 212

```

```

RESULT 15
AAM78582
ID AAM78582 standard; Protein: 605 AA.
XX
AC AAM78582;

```

```

XX 06-NOV-2001 (first entry)
DT
XX
DE Human protein SEQ ID NO 1244.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang J, Zhang D, Ren F, Chen R, Wang ZM;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR N-PSDB: AAK51715.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 3503-3504; 6221bp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 605 AA;

```

```

Query Match 100.0%; Score 240; DB 22; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALPARGDHIAENILSYLDAKSLCAELVCKEWRVTSQGLMKK 45
Db 183 ALPARGDHIAENILSYLDAKSLCAELVCKEWRVTSQGLMKK 227

```

```

Search completed: April 11, 2003, 11:48:12
Job time : 14.8672 secs

```

Fri Apr 11 13:31:07 2003

us-09-601-168b-2\_copy\_147\_191.rai

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 4.65895 Seconds  
(without alignments)  
284.191 Million cell updates/sec

Title: us-09-601-168b-2\_copy\_147\_191

Perfect score: 240  
Sequence: 1 ALPARGLDHIAENILSYLDA.....ALVCKEWRVYTSQMLMK 45

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCUTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/Backfill1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	237	98.8	517	1	US-08-190-802A-30 Sequence 30, Appl
2	237	98.8	517	4	US-08-477-346-30 Sequence 30, Appl
3	237	98.8	517	4	US-08-473-089-30 Sequence 30, Appl
4	237	98.8	517	4	US-08-487-072A-30 Sequence 30, Appl
5	226	94.2	42	4	US-09-172-841-1 Sequence 1, Appl
6	91	37.9	640	4	US-09-177-165A-30 Sequence 30, Appl
7	77	32.1	519	4	US-09-172-841-55 Sequence 37, Appl
8	73	30.4	38	4	US-09-172-841-37 Sequence 4, Appl
9	72.5	30.2	435	2	US-08-531-439B-4 Sequence 13, Appl
10	69	28.7	38	4	US-09-172-841-13 Sequence 29, Appl
11	68	28.3	33	4	US-09-172-841-29 Sequence 39, Appl
12	67	27.9	35	4	US-09-172-841-19 Sequence 19, Appl
13	65.5	27.3	35	4	US-09-172-841-33 Sequence 15, Appl
14	65	27.1	38	4	US-09-172-841-15 Sequence 416, Appl
15	64	26.7	38	4	US-08-905-223-416 Sequence 21, Appl
16	64	26.7	78	4	US-08-172-841-21 Sequence 51, Appl
17	63.5	26.5	35	4	US-09-172-841-51 Sequence 3, Appl
18	61	25.4	456	4	US-09-172-841-33 Sequence 47, Appl
19	58	24.2	38	4	US-09-172-841-47 Sequence 27, Appl
20	57	23.8	38	4	US-08-944-483-46 Sequence 46, Appl
21	57	23.8	234	4	US-09-172-841-27 Sequence 15, Appl
22	56	23.3	40	4	US-09-166-350-15 Sequence 23, Appl
23	53.5	22.3	321	4	US-09-172-841-23 Sequence 41, Appl
24	52	21.7	38	4	US-09-172-841-41 Sequence 12, Appl
25	51.5	21.5	43	4	US-09-173-151A-12 Sequence 147, Appl
26	51	21.2	182	1	US-08-127-954-147
27	51	21.2	182	1	US-08-127-954-147

28	51	21.2	216	4	US-08-914-375C-19 Sequence 19, Appl
29	51	21.2	324	1	US-08-597-236-10 Sequence 10, Appl
30	51	21.2	406	4	US-08-746-682A-10 Sequence 4, Appl
31	51	21.2	406	4	US-09-171-699A-2 Sequence 2, Appl
32	51	21.2	406	5	PCT-US94-02107-2 Sequence 14, Appl
33	51	21.2	610	4	US-09-173-151A-14 Sequence 16, Appl
34	51	21.2	614	4	US-09-173-151A-16 Sequence 16, Appl
35	51	21.2	701	2	US-08-533-669A-16 Sequence 16, Appl
36	51	21.2	701	4	US-09-183-861-16 Sequence 14, Appl
37	51	21.2	701	4	US-09-022-765-16 Sequence 14, Appl
38	50.5	21.0	106	2	US-08-820-825-14 Sequence 14, Appl
39	50.5	21.0	106	4	US-09-307-817-14 Sequence 14, Appl
40	50.5	21.0	106	4	US-09-734-036-14 Sequence 2, Appl
41	50.5	21.0	107	1	US-08-409-731A-2 Sequence 2, Appl
42	50.5	21.0	107	2	US-08-470-298B-2 Sequence 2, Appl
43	50.5	21.0	107	2	US-09-023-073A-2 Sequence 2, Appl
44	50.5	21.0	107	4	US-09-361-737-2 Sequence 57, Appl
45	50.5	20.8	39	4	US-09-172-841-57

ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
Sequence 30, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fahlao, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-190-802A-30

Query Match 98.8%, Score 237, DB 1, Length 517;  
Best local similarity 97.8%, Pred. No. 4.3e-24;  
Matches 44; Conservative 1; Mismatches 0; Gaps 0;

Db 112 ALPARGDHAENILSYLDAKSLCSAEVCKEYRVTSDGMLMK 156

## RESULT 2

US-08-477-346-30

Sequence 30, Application US/08477346

Patent No. 6262023

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theeol

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSER: Morrison &amp; Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,346

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,072

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13

US-08-477-346-30

Query Match 98.8%; Score 237; DB 4; Length 517;

Best Local Similarity 97.8%; Pred. No. 4.3e-24;

Matches 44; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPARGDHAENILSYLDAKSLCSAEVCKEYRVTSDGMLMK 45  
Db 112 ALPARGDHAENILSYLDAKSLCSAEVCKEYRVTSDGMLMK 156

RESULT 3

US-08-473-089-30

Sequence 30, Application US/08473089

Patent No. 6342368

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theeol

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSER: Morrison &amp; Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,089

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.22

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13

US-08-473-089-30

Query Match 98.8%; Score 237; DB 4; Length 517;

Best Local Similarity 97.8%; Pred. No. 4.3e-24;

Matches 44; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPARGDHAENILSYLDAKSLCSAEVCKEYRVTSDGMLMK 45  
Db 112 ALPARGDHAENILSYLDAKSLCSAEVCKEYRVTSDGMLMK 156

RESULT 4

US-08-487-072A-30

Sequence 30, Application US/08487072A

Patent No. 6423684

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theeol

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSER: Morrison &amp; Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:



```

; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0753
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-487-072A-30

Query Match          98.8%; Score 237; DB 4; Length 517;
Best Local Similarity 97.8%; Pred. No. 4,3e-24;
Matches 44; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALPARGLDHIAENILSYDAKSLCAELVCKEYRVTSQGLMK 45
DB 112 ALPARGLDHIAENILSYDAKSLCAELVCKEYRVTSQGLMK 156

RESULT 5
US-09-172-841-1
; Sequence 1, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-172-841-1

Query Match          94.2%; Score 226; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 6,6e-24;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPARGLDHIAENILSYDAKSLCAELVCKEYRVTSQGLMK 43
DB 1 LPARGLDHIAENILSYDAKSLCAELVCKEYRVTSQGLMK 42

RESULT 6
US-09-177-165A-30
; Sequence 30, Application US/09177165A
; Patent No. 6426205
; GENERAL INFORMATION:
; APPLICANT: Tyers, Mike
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
; FILE REFERENCE: 11757.10US01
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: US/09/177,165A
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/092,443
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 640

```

```

; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-177-165A-30

Query Match          37.9%; Score 91; DB 4; Length 640;
Best Local Similarity 50.0%; Pred. No. 0.00033;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 14 ILSTYDAKSLCAELVCKEYRVTSQGLMK 43
DB 195 ILSTYDCQSLCNATVCRKMKLADDDRW 224

RESULT 7
US-09-172-841-55
; Sequence 55, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-172-841-55

Query Match          32.1%; Score 77; DB 4; Length 519;
Best Local Similarity 34.9%; Pred. No. 0.02;
Matches 15; Conservative 8; Mismatches 16; Indels 4; Gaps 1;

OY 6 GLDHIAE----NLSTYDAKSLCAELVCKEYRVTSQGLMK 44
DB 32 GITHLPPEVWLSTFYLNPQLRCQSVSMKNSQLTKTSLMK 74

RESULT 8
US-09-172-841-37
; Sequence 37, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-172-841-37

Query Match          30.4%; Score 73; DB 4; Length 38;
Best Local Similarity 53.3%; Pred. No. 0.003;
Matches 16; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

OY 14 ILSTYDAKSLCAELVCKEYRVTSQGLMK 43
DB 9 ILSTYDAVSLIQAAQVNRKNNMELASDVLW 38

```

RESULT 9  
US-08-531-439B-4  
; Sequence 4, Application US/08531439B  
; Patent No. 5981702  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Hui  
; APPLICANT: Beach, David  
; TITLE OF INVENTION: Cyclin/CDK Associated Proteins,  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Elliot  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/531,439B  
; FILING DATE: 21-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: CSV-006.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 435 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-531-439B-4

Query Match  
Best Local Similarity 30.2%; Score 72.5; DB 2; Length 435;  
Matches 14; Conservative 2; Mismatches 6; Indels 5; Gaps 1;

QY 23 LCAAEI---VCKEMYRTSDGMLWK 44  
DB 124 LCLPELKVSGCKRWYLADESLWQ 150

RESULT 10  
US-09-172-841-13  
; Sequence 13, Application US/09172841  
; Patent No. 6232081  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeffrey W.  
; APPLICANT: Ellledge, Stephen J.  
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES  
; FILE REFERENCE: BCM-03510  
; CURRENT APPLICATION NUMBER: US/09/172,841  
; EARLIER FILING DATE: 1998-10-15  
; EARLIER APPLICATION NUMBER: 08/951,621  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 38  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-172-841-13

Query Match  
Best Local Similarity 28.7%; Score 69; DB 4; Length 38;  
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Matches 12; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 13 NILSYLDKSAIAELVCKEMYRTSDGMLW 43  
DB 8 SIFSILNPDELRCRSGVSMKMSLTKTGSIW 38

RESULT 11  
US-09-172-841-29  
; Sequence 29, Application US/09172841  
; Patent No. 6232081  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeffrey W.  
; APPLICANT: Ellledge, Stephen J.  
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES  
; FILE REFERENCE: BCM-03510  
; CURRENT APPLICATION NUMBER: US/09/172,841  
; EARLIER FILING DATE: 1998-10-15  
; EARLIER APPLICATION NUMBER: 08/951,621  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-09-172-841-29

Query Match  
Best Local Similarity 28.3%; Score 68; DB 4; Length 33;  
Matches 15; Conservative 2; Mismatches 11; Indels 4; Gaps 1;

QY 2 LPARGLDHIAENILSYLDKSAIAELVCKEW 33  
DB 1 LPA---LITFKIFSOLDIRSRASLTGCRGW 28

RESULT 12  
US-09-172-841-39  
; Sequence 39, Application US/09172841  
; Patent No. 6232081  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeffrey W.  
; APPLICANT: Ellledge, Stephen J.  
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES  
; FILE REFERENCE: BCM-03510  
; CURRENT APPLICATION NUMBER: US/09/172,841  
; EARLIER FILING DATE: 1998-10-15  
; EARLIER APPLICATION NUMBER: 08/951,621  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 39  
; LENGTH: 38  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-09-172-841-39

Query Match  
Best Local Similarity 27.9%; Score 67; DB 4; Length 38;  
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 10 IAEHILSYLDKSAIAELVCKEMYRTSDGMLW 43  
DB 5 ILVKILSYLDAVTLVCIQVSRFRYHLADNDLIW 38

RESULT 13  
US-09-172-841-19  
; Sequence 19, Application US/09172841  
; Patent No. 6232081  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeffrey W.

```

APPLICANT: Ellledge, Stephen J.
FILE REFERENCE: BCM-03510
CURRENT APPLICATION NUMBER: US/09/172,841
CURRENT FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: 08/951,621
EARLIER FILING DATE: 1997-10-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 19
LENGTH: 35
TYPE: PRT
ORGANISM: Homo sapiens
US-09-172-841-19

```

```

Query Match 27.3%; Score 65.5; DB 4; Length 35r
Best Local Similarity 46.7%; Pred. No. 0.029; 9; Indels 3; Gaps 1;
Matches 14; Conservative 4; Mismatches 9;

```

```

QY 14 ILISYDAKSLCAELVCKEYRVTSQGLM 43
DB 9 ILISYLAATDCLASCY---MODLANDELLM 35

```

```

RESULT 14
US-09-172-841-33
Sequence 33; Application US/09172841
Patent No. 6232081
GENERAL INFORMATION:
APPLICANT: Harper, Jeffrey W.
APPLICANT: Ellledge, Stephen J.
TITLE OF INVENTION: F-BOX PROTEINS AND GENES
FILE REFERENCE: BCM-03510
CURRENT APPLICATION NUMBER: US/09/172,841
CURRENT FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: 08/951,621
EARLIER FILING DATE: 1997-10-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 33
LENGTH: 38
TYPE: PRT
ORGANISM: Mus musculus
US-09-172-841-33

```

```

Query Match 27.1%; Score 65; DB 4; Length 38;
Best Local Similarity 36.1%; Pred. No. 0.037; 18; Indels 0; Gaps 0;
Matches 13; Conservative 5; Mismatches 18;

```

```

QY 8 DHTAENILSYLDAKSLCAELVCKEYRVTSQGLM 43
DB 3 DSIYVQIFILSGPADVLAAGLVCRQWQAVSRDEFLM 38

```

```

RESULT 15
US-09-172-841-15
Sequence 15; Application US/09172841
Patent No. 6232081
GENERAL INFORMATION:
APPLICANT: Harper, Jeffrey W.
APPLICANT: Ellledge, Stephen J.
TITLE OF INVENTION: F-BOX PROTEINS AND GENES
FILE REFERENCE: BCM-03510
CURRENT APPLICATION NUMBER: US/09/172,841
CURRENT FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: 08/951,621
EARLIER FILING DATE: 1997-10-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 15
LENGTH: 38
TYPE: PRT
ORGANISM: Mus musculus

```

```

US-09-172-841-15
Query Match 26.7%; Score 64; DB 4; Length 38;
Best Local Similarity 35.5%; Pred. No. 0.05; 14; Indels 0; Gaps 0;
Matches 11; Conservative 6; Mismatches 14;

```

```

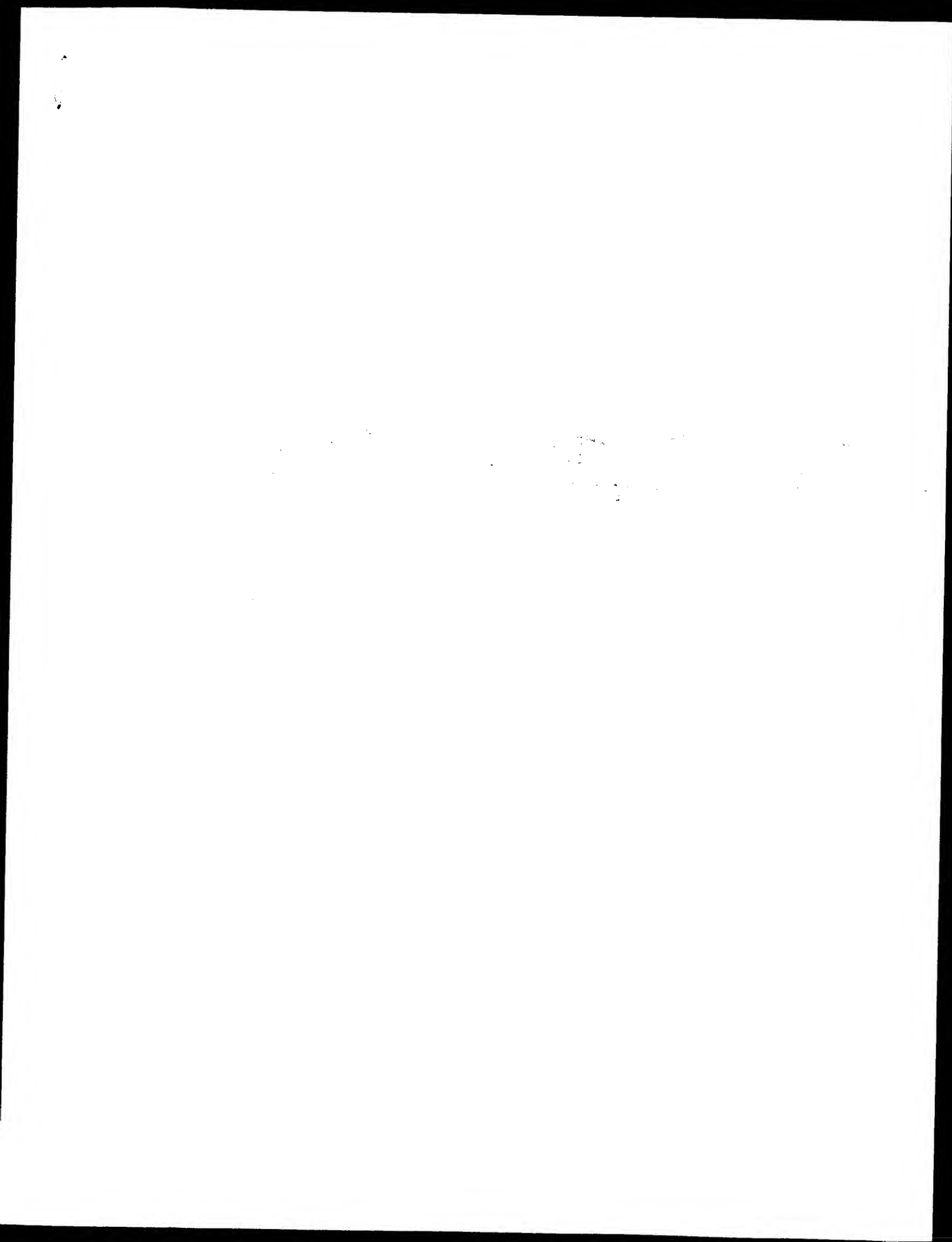
QY 13 NILSYDAKSLCAELVCKEYRVTSQGLM 43
DB 8 SIFSILNPDELRCRCQSVSTKWSQIAKTGSLM 38

```

```

Search completed: April 10, 2003, 12:51:17
Job time : 6.65895 secs

```



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:49:06 ; Search time 4.60414 Seconds  
(Without alignments)  
597.532 Million cell updates/sec

Title: US-09-601-168B-2\_COPY\_147\_191

Perfect score: 240

Sequence: 1 ALPARGLDHIAENILSYLDA.....AEVCKEYRVTSIDGMLMK 45

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEM\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCIT\_NEM\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEM\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEM\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCITUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US09\_NEM\_PUB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US10\_NEM\_PUB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	100.0	219	9 US-10-023-530-2	Sequence 2, Appli
2	240	100.0	569	9 US-10-038-010-8	Sequence 8, Appli
3	240	100.0	569	12 US-10-042-417-2	Sequence 2, Appli
4	231	96.2	43	12 US-10-042-417-15	Sequence 15, Appli
5	210	87.5	265	9 US-10-116-016-30	Sequence 30, Appli
6	210	87.5	265	10 US-09-764-848-30	Sequence 30, Appli
7	91	37.9	9	US-10-060-019-30	Sequence 30, Appli
8	83	34.6	39	12 US-10-042-417-71	Sequence 71, Appli
9	83	34.6	483	12 US-10-042-417-48	Sequence 48, Appli
10	79	32.9	678	10 US-09-801-368-314	Sequence 314, App
11	78	32.5	197	12 US-10-042-417-32	Sequence 32, Appli
12	77	32.1	590	12 US-10-042-417-52	Sequence 52, Appli
13	77	31.7	39	12 US-10-042-417-64	Sequence 64, Appli
14	74	30.8	39	12 US-10-042-417-73	Sequence 73, Appli
15	72.5	30.2	39	12 US-10-042-417-22	Sequence 22, Appli
16	70	29.2	540	9 US-09-213-888-7	Sequence 7, Appli
17	70	29.2	540	9 US-09-213-888-10	Sequence 10, Appli
18	70	29.2	540	9 US-09-328-877A-7	Sequence 7, Appli
19	70	29.2	540	9 US-09-328-877A-10	Sequence 10, Appli

20	70	29.2	545	9 US-09-213-888-6	Sequence 6, Appli
21	70	29.2	545	9 US-09-328-877A-6	Sequence 6, Appli
22	70	29.2	553	9 US-09-213-888-5	Sequence 5, Appli
23	70	29.2	553	9 US-09-328-877A-5	Sequence 5, Appli
24	70	29.2	559	9 US-09-213-888-9	Sequence 9, Appli
25	70	29.2	559	9 US-09-328-877A-9	Sequence 9, Appli
26	70	29.2	589	9 US-09-213-888-8	Sequence 8, Appli
27	70	29.2	589	9 US-09-328-877A-8	Sequence 8, Appli
28	70	29.2	592	9 US-09-213-888-4	Sequence 4, Appli
29	70	29.2	592	9 US-09-328-877A-4	Sequence 4, Appli
30	70	29.2	626	9 US-09-213-888-21	Sequence 21, Appli
31	70	29.2	626	9 US-09-328-877A-21	Sequence 21, Appli
32	70	29.2	627	9 US-09-213-888-3	Sequence 3, Appli
33	70	29.2	627	9 US-09-328-877A-3	Sequence 3, Appli
34	70	29.2	666	9 US-09-213-888-27	Sequence 27, Appli
35	70	29.2	666	9 US-09-328-877A-27	Sequence 27, Appli
36	70	29.2	669	9 US-09-213-888-25	Sequence 25, Appli
37	70	29.2	669	9 US-09-328-877A-25	Sequence 25, Appli
38	66.5	27.7	36	12 US-10-042-417-61	Sequence 61, Appli
39	66.5	27.7	319	9 US-10-028-072-134	Sequence 134, App
40	66.5	27.7	319	9 US-10-121-049-134	Sequence 134, App
41	66.5	27.7	319	9 US-10-123-904-134	Sequence 134, App
42	66.5	27.7	319	9 US-10-140-470-134	Sequence 134, App
43	66.5	27.7	319	9 US-10-175-746-134	Sequence 134, App
44	66.5	27.7	319	9 US-10-176-918-134	Sequence 134, App
45	66.5	27.7	319	9 US-10-176-921-134	Sequence 134, App

# ALIGNMENTS

RESULT 1  
US-10-023-530-2  
Sequence 2, Application US/10023530  
Publication No. US20030007956A1  
GENERAL INFORMATION:  
APPLICANT: LEBERAIN, Pierre  
APPLICANT: BENAROUS, Richard  
APPLICANT: BIOT, Guillaume  
TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP  
FILE REFERENCE: B4717A  
CURRENT APPLICATION NUMBER: US/10/023,530  
CURRENT FILING DATE: 2002-04-22  
PRIOR APPLICATION NUMBER: 60/256,276  
PRIOR FILING DATE: 2000-12-18  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 219  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: BetaTRCP  
LOCATION: (1)..(219)  
OTHER INFORMATION: F-box protein  
US-10-023-530-2

Query Match 100.0%; Score 240; DB 9; Length 219;  
Best Local Similarity 100.0%; Pred. No. 2.2e-25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 45  
DB 147 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 191  
RESULT 2  
US-10-038-010-8  
Sequence 8, Application US/10038010  
Publication No. US20030040089A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS

```
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: beta-Trip1
; LOCATION: (1)..(569)
; OTHER INFORMATION:
US-10-038-010-8
```

```
Query Match
Best local Similarity 100.0%; Score 240; DB 9; Length 569;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 45
DB 147 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 191
```

```
RESULT 3
US-10-042-417-2
; Sequence 2, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-2
```

```
Query Match
Best local Similarity 100.0%; Score 240; DB 12; Length 569;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 45
DB 147 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 191
```

```
RESULT 4
US-10-042-417-15
; Sequence 15, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 15
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-15
```

```
Query Match
Best local Similarity 96.2%; Score 231; DB 12; Length 43;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 LPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 44
DB 1 LPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 43
```

```
RESULT 5
US-10-116-016-30
; Sequence 30, Application US/10116016
; Publication No. US20030054379A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT208C1
; CURRENT APPLICATION NUMBER: US/10/116,016
; PRIOR APPLICATION REMOVED - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-016-30
```

```
Query Match
Best local Similarity 87.5%; Score 210; DB 9; Length 265;
Matches 39; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 45
DB 115 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 159
```

```
RESULT 6
US-09-764-848-30
; Sequence 30, Application US/09764848
; Patent No. US20020077270A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT208
; CURRENT APPLICATION NUMBER: US/09/764,848
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-848-30
```

```
Query Match
Best local Similarity 87.5%; Score 210; DB 10; Length 265;
Matches 39; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 45
DB 115 ALPARGDHTAENILSYLDKSLCAELVCKEYRVTSDGMLMK 159
```

```
RESULT 7
US-10-060-019-30
```

```
Sequence 30, Application US/10060019
Publication No. US20030003564A1
GENERAL INFORMATION:
APPLICANT: Tyers, Mike
APPLICANT: Willems, Andrew
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBQUITIN
TITLE OF INVENTION: DEPENDENT PROTEOLYSIS
FILE REFERENCE: 11757.10USU1
CURRENT APPLICATION NUMBER: US/10/060,019
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US/09/177,165
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/092,443
PRIOR FILING DATE: 1998-07-10
PRIOR APPLICATION NUMBER: 60/063,254
PRIOR FILING DATE: 1997-10-24
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 30
LENGTH: 640
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-060-019-30

Query Match          37.9%; Score 91; DB 9; Length 640;
Best Local Similarity 50.0%; Pred. No. 0.00023;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 14 ILSTYDAKSLCAAEIYCKEYRYVTSQGLMK 43
Db 195 ILSTYDCOSLCNATRVCKKMKLADDDRW 224

RESULT 8
US-10-042-417-71
Sequence 71, Application US/10042417
Patent No. US20020123082A1
GENERAL INFORMATION:
APPLICANT: Pagano, M.
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
FILE REFERENCE: 5914-090-999
CURRENT APPLICATION NUMBER: US/10/042,417
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: 60/260,179
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 71
LENGTH: 39
TYPE: PRT
ORGANISM: Homo sapiens
US-10-042-417-71

Query Match          34.6%; Score 83; DB 12; Length 39;
Best Local Similarity 37.8%; Pred. No. 0.00012;
Matches 14; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 8 DHTAENILSTYDAKSLCAAEIYCKEYRYVTSQGLMK 44
Db 3 DHTSWQIFSLPTNQLRCARVCRRWYNLAMPRLMK 39

RESULT 9
US-10-042-417-48
Sequence 48, Application US/10042417
Patent No. US20020123082A1
GENERAL INFORMATION:
APPLICANT: Pagano, M.
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
FILE REFERENCE: 5914-090-999
CURRENT APPLICATION NUMBER: US/10/042,417
```

```
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: 60/260,179
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 48
LENGTH: 483
TYPE: PRT
ORGANISM: Homo sapiens
US-10-042-417-48

Query Match          34.6%; Score 83; DB 12; Length 483;
Best Local Similarity 37.8%; Pred. No. 0.0021;
Matches 14; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 8 DHTAENILSTYDAKSLCAAEIYCKEYRYVTSQGLMK 44
Db 111 DHTSWQIFSLPTNQLRCARVCRRWYNLAMPRLMK 147

RESULT 10
US-09-801-368-314
Sequence 314, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Call, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 314
LENGTH: 678
TYPE: PRT
ORGANISM: Aspergillus nidulans
US-09-801-368-314

Query Match          32.9%; Score 79; DB 10; Length 678;
Best Local Similarity 40.0%; Pred. No. 0.011;
Matches 18; Conservative 5; Mismatches 18; Indels 4; Gaps 1;

QY 1 ALPARGLDHTAENILSTYDAKSLCAAEIYCKEYRYVTSQGLMK 45
Db 183 ALP-----PEIAFKILCYLDTTSLCKASQVSRGWRALADDDVYVWHR 223

RESULT 11
US-10-042-417-32
Sequence 32, Application US/10042417
Patent No. US20020123082A1
GENERAL INFORMATION:
APPLICANT: Pagano, M.
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
FILE REFERENCE: 5914-090-999
CURRENT APPLICATION NUMBER: US/10/042,417
```







—

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 5.20706 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_147\_191

Sequence: 1 ALPARGIDHIAENILSYLDA.....ALVCKEYRYVTSQGLMKK 45

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	237	98.8	518	2 B48088	beta-transducin re
2	91	37.9	640	2 S49932	MET30 protein - ye
3	88	36.7	605	2 T38932	probable sulfur me
4	81.5	34.0	662	2 T50319	hypothetical F-box
5	80.5	33.5	775	2 T45136	WD repeat protein
6	79	32.9	506	2 T50211	WD repeat protein
7	78	32.5	197	2 G88523	truncated hypotet
8	78	32.5	489	2 S44609	hypothetical prote
9	75.5	31.5	614	2 T09902	hypothetical prote
10	75	31.2	419	2 T32441	hypothetical prote
11	74	30.8	151	2 T09884	hypothetical prote
12	72.5	30.2	435	2 T39171	hypothetical prote
13	70.5	29.4	592	2 T52139	cyclin A/CDK2-asso
14	68.5	28.5	449	2 B85069	LRR-containing F-b
15	68	28.3	324	2 E84505	hypothetical prote
16	66	27.5	304	2 G85068	hypothetical prote
17	66	27.1	411	2 T47951	hypothetical prote
18	62	25.8	418	2 T16410	hypothetical prote
19	62	25.8	431	2 T06019	sulfur controller
20	62	25.8	650	2 T46660	hypothetical prote
21	61	25.4	382	2 T45851	hypothetical prote
22	61	25.4	934	2 T05201	hypothetical prote
23	60	25.0	276	2 T52349	F-box protein FBL2
24	59.5	24.8	463	2 T39987	probable regulator
25	59.5	24.8	577	2 F86261	FLK23.7 protein -
26	59	24.6	141	2 S30832	hypothetical prote
27	59	24.6	302	2 F83068	N7-like protein [i
28	59	24.6	307	2 A85069	N7-like protein [i
29	59	24.6	701	2 T16607	hypothetical prote

30	58.5	24.4	160	2 F86283	protein T1SD22.6 [
31	58.5	24.4	561	2 H86423	hypothetical prote
32	58	24.2	76	2 C43559	homeotic protein R
33	57.5	24.0	465	2 D96567	F6D8.13 (imported)
34	57	23.8	122	2 T08915	hypothetical prote
35	57	23.8	262	1 A31372	granzyme A (EC 3.4
36	57	23.8	269	2 T16487	hypothetical prote
37	57	23.8	279	2 G96545	hypothetical prote
38	57	23.8	384	2 G84482	hypothetical prote
39	57	23.8	384	2 T49128	hypothetical prote
40	57	23.8	554	2 AD1572	propanediol dehydr
41	57	23.8	554	2 A11218	hypothetical prote
42	56	23.3	475	2 T47778	hypothetical prote
43	55	22.9	188	2 C97438	conserved hypotet
44	55	22.9	188	2 AF2656	hypothetical prote
45	55	22.9	346	2 T45748	

## ALIGNMENTS

RESULT 1  
B48088  
beta-transducin repeat-containing protein - African clawed frog  
N:Alternate names: beta-Trop  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000  
R:Spavak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.  
Mol. Cell. Biol. 13, 4953-4966, 1993  
A>Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase  
A:Reference number: A48088; MID:93330289; PMID:8393141  
A:Accession: B48088  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-518 <SP>  
A:Cross-references: GB:M98268; NID:9295542; PID:NAA02810.1; PID:9295543  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: duplication  
F:431-462/Domain: WD repeat homology <WD1>

Query Match 98.8% Score 237; DB 2; Length 518;  
Best Local Similarity 97.8%; Pred. No. 2; 1e-23;  
Matches 44; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGIDHIAENILSYLDAKSLCAELVCKEYRYVTSQGLMKK 45  
DB 112 ALPARGIDHIAENILSYLDAKSLCAELVCKEYRYVTSQGLMKK 156

## RESULT 2

S49932  
MET30 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein Y19905.02; protein Y1L046w  
C:Species: Saccharomyces cerevisiae  
C>Date: 26-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 26-May-2000  
C:Accession: S49932; S43750  
R:Odell, C.; Bowman, S.  
Submitted to the EMBL Data Library, December 1994  
A:Reference number: S49931  
A:Accession: S49932  
A:Molecule type: DNA  
A:Residues: 1-640 <CODE>  
A:Cross-references: GB:247047; EMBL:246861; NID:9603997; PID:9763300; MIPS:Y1L046w  
R:Thomas, D.; Charest, H.; Barbey, R.; Surdin-Kerjan, Y.  
Submitted to the EMBL Data Library, December 1993  
A:Reference number: S43750  
A:Accession: S43750  
A:Molecule type: DNA  
A:Residues: 1-60, '1', 62-640 <THO>  
A:Cross-references: EMBL:L26505; NID:9432493; PID:9432494  
C:Genetics: SCD:MET30  
A:Gene: SCD:MET30

A:Cross-references: SGD:S0001308; MIPS:YIL046w  
A:Map position: 9L

C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
E:298-329/Domain: WD repeat homology <MD1>  
E:338-369/Domain: WD repeat homology <MD2>  
E:374-409/Domain: WD repeat homology <MD3>  
F:417-450/Domain: WD repeat homology <MD4>

Query Match  
Best Local Similarity 37.9%; Score 91; DB 2; Length 640;  
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 14 ILSTLDKSLCAAEIVCKEYRVTSQGMW 43  
Db 195 ILSTLDQSLCAATRVCKRWKRLADDDRW 224

#### RESULT 3

probable sulfur metabolite control protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 26-May-2000

C:Accession: T38932  
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, April 1997

A:Reference number: 221818  
A:Accession: T38932  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-605 <BAD>  
A:Cross-references: EMBL:294864; PIDN:CA808168.1; GSPDB:GN00066; SPDB:SPAC57A10.05C

A:Experimental source: strain 972h-; cosmid c57A10  
C:Genetics:  
A:Gene: SPDB:SPAC57A10.05C  
A:Map position: 1

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match  
Best Local Similarity 36.7%; Score 88; DB 2; Length 605;  
Matches 16; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 10 IAENILSYLDKSLCAAEIVCKEYRVTSQGMW 45  
Db 117 ISFRIISFLDARSLCAQAOVSKRWKRLADDDYVHR 152

#### RESULT 4

hypothetical F-box domain protein [imported] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000

C:Accession: T50319  
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; Lelaure, V.; Gallibert, F.  
submitted to the EMBL Data Library, January 2000

A:Reference number: 225061  
A:Accession: T50319  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-662 <MCD>  
A:Cross-references: EMBL:AL136536; PIDN:CA866450.1; GSPDB:GN00067; SPDB:SPAC1703.06

A:Experimental source: strain 972h(-); cosmid c1703  
C:Genetics:  
A:Gene: SPDB:SPAC1703.06  
A:Map position: 2

A:Introns: 37/1

Query Match  
Best Local Similarity 34.0%; Score 81.5; DB 2; Length 662;  
Matches 16; Conservative 9; Mismatches 12; Indels 15; Gaps 1;

QY 8 DHIATN-----ILSYLDKSLCAAEIVCKEYRVTSQGMW 44  
Db 21 DHSSNNTNRTVNLUPKEIIILIFSLDPRSLLSACCTCKYMKRLSDLSMR 72

#### RESULT 5

WD repeat protein popl [imported] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T45136; T40157  
R:Kominami, K.; Toda, T.  
submitted to the EMBL Data Library, September 1996

A:Description: Fission yeast WD repeat protein Popl is involved in maintenance of plo  
A:Reference number: 223925  
A:Accession: T45136

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-775 <KOM>

A:Cross-references: EMBL:Y08391; PIDN:CA69671.1  
A:Experimental source: strain h- 972

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.  
submitted to the EMBL Data Library, March 1998  
A:Reference number: 221842  
A:Accession: T40157

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-775 <WOOD>

A:Cross-references: EMBL:AL022103; PIDN:CA17898.1; GSPDB:GN00067; SPDB:SPAC2G2.18  
A:Experimental source: strain 972h-; cosmid c2G2  
C:Genetics:  
A:Gene: SPAC2G2.18  
A:Map position: 2

A:Note: popl+

Query Match  
Best Local Similarity 33.5%; Score 80.5; DB 2; Length 775;  
Matches 17; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 10 IAENILSYLDKSLCAAEIVCKEYRVTSQGMW 44  
Db 308 ITNLVLTFLDAPSLCAVSOVSHMYKLVSSNEELW 343

#### RESULT 6

WD repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000

C:Accession: T50211  
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.;  
submitted to the EMBL Data Library, January 2000

A:Reference number: 225046  
A:Accession: T50211

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-506 <MCD>

A:Cross-references: EMBL:AL136538; PIDN:CA866464.1; GSPDB:GN00066; SPDB:SPAC30.05  
A:Experimental source: strain 972h(-); cosmid c30  
C:Genetics:  
A:Gene: SPAC2956.01; SPDB:SPAC30.05  
A:Map position: 1

A:Introns: 43/1; 74/3  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match  
Best Local Similarity 32.9%; Score 79; DB 2; Length 506;  
Matches 13; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 8 DHIATNLSYLDKSLCAAEIVCKEYRVTSQGMW 44  
Db 78 BEVSLRVFSYLDQDLCKKMLSKRWKRLLEDPGIWK 114

#### RESULT 7

G88523

1

1



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:51:26 ; Search time 3.2386.Seconds  
(without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_147\_191  
Perfect score: 240  
Sequence: 1 ALPARGDHTAENILSTYLA.....AEVCKEWRVTSQGLMKK 45

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	240	100.0	605	FWIA_HUMAN	09y297 homo sapien
2	237	98.8	518	TRCB_XENLA	091854 xenopus lae
3	210	87.5	542	FW1B_HUMAN	09akb1 homo sapien
4	91	37.9	640	MT30_YEAST	P39014 saccharomyc
5	88	36.7	605	POPI_SCHPO	P87053 schizosacch
6	81.5	34.0	662	POFA_SCHPO	09y744 schizosacch
7	80.5	33.5	775	POPI_SCHPO	P87060 schizosacch
8	79	32.9	506	POFB_SCHPO	009855 schizosacch
9	79	32.9	461	SCOB_EMENT	000659 emeritella
10	78	32.5	684	YKK7_CAEEL	P34284 caenorhabdi
11	77	32.1	684	CC4_CANAL	P53699 candida alb
12	66.5	27.7	319	FBX8_HUMAN	09and0 homo sapien
13	64.5	26.9	419	FBX8_MOUSE	09qzn3 mus musculu
14	64	26.7	412	SCO2_NEUCR	P57775 homo sapien
15	62	25.8	650	FBM4_MOUSE	001277 neurospora
16	61	25.4	410	FBM4_MOUSE	09jmf2 mus musculu
17	59	24.6	141	YEFS_YEAST	P32616 saccharomyc
18	59	24.6	522	FBX7_HUMAN	09y311 homo sapien
19	59	24.6	551	FX10_HUMAN	09uk96 homo sapien
20	59	24.6	551	REFR_SPYKA	008517 swinepox vi
21	59	24.6	665	LI23_CAEEL	009990 caenorhabdi
22	58	24.2	76	HXCA_RAT	P18665 rattus norv
23	57	23.8	262	GRAA_HUMAN	P12544 homo sapien
24	55	22.9	422	FBM2_HUMAN	09uk88 homo sapien
25	53.5	22.3	447	FBX9_HUMAN	09uk97 homo sapien
26	53.5	22.3	500	YDAK_YEAST	P28817 saccharomyc
27	53.5	22.3	593	MTF2_HUMAN	09y483 homo sapien
28	53	22.1	318	FX24_HUMAN	075426 homo sapien
29	52	21.7	263	TRUA_BACHD	09z910 bacillus ha
30	52	21.7	366	XYLB_PSEPU	P39849 pseudomonas
31	52	21.7	447	FBX2_HUMAN	09uk15 homo sapien
32	51.5	21.5	295	FBX2_HUMAN	09uk42 homo sapien
33	51.5	21.5	332	ILVC_ARCFU	O28294 archaeoglob

34	51	21.2	491	1	VIE1_HCMVA	P13202 human cytom
35	51	21.2	491	1	VIE1_HCMVT	P03169 human cytom
36	51	21.2	701	1	HSB3_LETAM	P27741 leishmania
37	51	21.2	727	1	VR22_CAEEL	009639 caenorhabdi
38	50	20.8	242	1	YB08_PSEAE	09hy11 pseudomonas
39	50	20.8	382	1	V382_ASEFL5	P26712 african swi
40	50	20.8	493	1	V100_STRPN	097p44 streptococc
41	50	20.8	664	1	TXB7_CAEEL	023467 caenorhabdi
42	49.5	20.6	517	1	VGLG_VSVJO	P04882 vesicular s
43	49	20.4	666	1	YMI8_YEAST	004511 saccharomyc
44	48.5	20.2	134	1	RET5_HUMAN	P82980 homo sapien
45	48.5	20.2	254	1	ADH_DROCK	000670 drosophila

## ALIGNMENTS

RESULT 1	FWIA_HUMAN	STANDARD:	PRT:	605 AA.
ID	FWIA_HUMAN			
AC	09y297; 09y213:			
DT	16-OCT-2001 (rel. 40, Created)			
DT	16-OCT-2001 (rel. 40, Last sequence update)			
DT	16-OCT-2001 (rel. 40, Last annotation update)			
DE	F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TRCP)			
DE	(E3RSIKappab) (pikappabalpha-E3 receptor subunit).			
GN	FBXWIA OR FBWIA OR FBXW1 OR FBXW2 OR FBXW3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	MEDLINE=90075339; PubMed=9859996;			
RX	Varon A., Hatzubal A., Davis M., Lavan I., Amit S., Manning A.M.,			
RA	Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;			
RT	"Identification of the receptor component of the Ikappabalpha-			
RL	ubiquitin ligase."			
RL	Nature 396:590-594(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RP	TISSUE=Lymphoid;			
KC	MEDLINE=98325370; PubMed=9660940;			
RA	Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,			
RA	Thomas D., Strebel K., Benarous R.;			
RT	"A novel human WD protein, h-beta TRCP, that interacts with HIV-1 Vpu			
RT	connects CD4 to the ER degradation pathway through an F-box motif."			
RL	Mol. Cell 1:565-574(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RP	MEDLINE=20003060; PubMed=10531035;			
RX	Cenciarelli C., Chlaud D.S., Guardavaccaro D., Parks W., Vidal M.,			
RA	Pagano M.;			
RT	"Identification of a family of human F-box proteins."			
RL	Curr. Biol. 9:1177-1179(1999).			
RN	[4]			
RP	CHARACTERIZATION.			
RP	MEDLINE=99145464; PubMed=990852;			
RX	Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,			
RA	Harper J.W.;			
RT	"The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically			
RT	with phosphorylated destruction motifs in I-kappa-B-alpha and			
RT	beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro."			
RL	Genes Dev. 13:270-283(1999).			
CC	- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA			
CC	(PIKAPPABALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR			
CC	UBIQUITINATION AND DEGRADATION.			
CC	- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.			
CC	- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	- FUNCTIONAL PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2: ARE			
CC	PRODUCED BY ALTERNATIVE SPLICING.			
CC	- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.			
CC	- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).			

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: AF101784; AAD08702.1; -
DR EMBL: Y14153; CA74572.1; -
DR EMBL: AF129530; AAF04464.1; -
DR Genew: HGNC:1144; BTRC.
DR MIM: 603482; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR Prodom: PD000018; WD40; 4.
DR SMART: SM00320; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS0181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 6.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.
KW DOMAIN 190 228 F-Box.
FT REPEAT 301 338 WD 1.
FT REPEAT 341 378 WD 2.
FT REPEAT 381 418 WD 3.
FT REPEAT 424 461 WD 4.
FT REPEAT 464 503 WD 5.
FT REPEAT 505 541 WD 6.
FT REPEAT 553 590 WD 7.
FT VARSPLIC 17 52 MISSING (IN ISOPFORM 2).
SQ SEQUENCE 605 AA; 68866 MW; 466793B7E400FD37 CRC64;

Query Match
Best Local Similarity 100.0%; Score 240; DB 1; Length 605;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIAENILSYLDAKSLCAELVCKEYRYTSDGMLMKK 45
Db 183 ALPARGLDHIAENILSYLDAKSLCAELVCKEYRYTSDGMLMKK 227

RESULT 2
TROB_XENLA STANDARD; PRT; 518 AA.
AC Q91854; P70037; P70038;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-TrCP (beta-transducin repeat-containing protein).
GN FBXW1 OR BTRCP.
OS Xenopus laevis (African Clawed Frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE=9330289; PubMed=8393141;
RA Spevak W., Kelper B.D., Stratowa C., Castanon M.J.;
"Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
anaphase are rescued by Xenopus CDNA's encoding N-ras or a protein
with beta-transducin repeats.";
RL Mol. Cell. Biol. 13:4953-4966(1993).
RN 12
RP SEQUENCE OF 302-518 FROM N.A.
RA MEDLINE=97109804; PubMed=8952061;
RA Hudson J.W., Alarcon V.B., Elinson R.P.;

```

```

RP "identification of new localized RNAs in the Xenopus oocyte by
RT differential display PCR.";
RL Dev. Genet. 19:190-198(1996).
CC -1- FUNCTION: probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
CC TADPOLE EMBRYO.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: M98268; AAA02810.1; -
DR EMBL: U63921; AAB49671.1; -
DR EMBL: U63922; AAB49672.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR Prodom: PD000018; WD40; 4.
DR SMART: SM00320; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS0181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 6.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Ubl conjugation pathway; Repeat; WD repeat.
KW DOMAIN 119 157 F-Box.
FT REPEAT 230 258 WD 1.
FT REPEAT 270 298 WD 2.
FT REPEAT 310 338 WD 3.
FT REPEAT 353 381 WD 4.
FT REPEAT 393 421 WD 5.
FT REPEAT 433 461 WD 6.
FT REPEAT 482 510 WD 7.
FT REPEAT 502 530 GFM -> EFR (IN REF. 2).
FT CONFLICT 302 516 GFM -> AAH (IN REF. 2).
FT CONFLICT 516 518 GLA -> AAH (IN REF. 2).
SQ SEQUENCE 518 AA; 59507 MW; 2A52EC19028127F3 CRC64;

Query Match
Best Local Similarity 98.8%; Score 237; DB 1; Length 518;
Matches 44; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIAENILSYLDAKSLCAELVCKEYRYTSDGMLMKK 45
Db 112 ALPARGLDHIAENILSYLDAKSLCAELVCKEYRYTSDGMLMKK 156

RESULT 3
FWIB_HUMAN STANDARD; PRT; 542 AA.
AC Q9UKB1; Q9YAC6; Q9P2S8; Q9P2S9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TrCP2).
GN FBXW1B OR FWIB OR BTRCP2 OR KIA00696.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```



NCBI\_TaxID=9606;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=20003600; PubMed=10531035;  
 RA Cenciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Fetal lung;  
 RX MEDLINE=20160458; PubMed=10694485;  
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,  
 RA Katoh M.;  
 RT "Molecular cloning and genomic structure of the betatropin gene on  
 RT chromosome 5q35.1.";  
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potentially).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A, B AND C (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF176022; AAF04528.1; -;  
 DR EMBL: AB033279; BAA92329.1; -;  
 DR EMBL: AB033280; BAA92330.1; -;  
 DR EMBL: AB033281; BAA92331.1; -;  
 DR EMBL: AB014596; BAA31671.1; ALT\_INIT.  
 DR Genew: HGNC:13607; FBXW1B.  
 DR MIM: 605651; -;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 FT DOMAIN 129 167 F-BOX.  
 FT REPEAT 238 275 WD 1.  
 FT REPEAT 278 315 WD 2.  
 FT REPEAT 318 355 WD 3.  
 FT REPEAT 355 398 WD 4.  
 FT REPEAT 398 440 WD 5.  
 FT REPEAT 440 478 WD 6.  
 FT REPEAT 478 527 WD 7.  
 FT REPEAT 527 564 WD 8.  
 FT VARSPLIC 16 49 MISSING (IN ISOFORM A).

FT VARSPLIC 16 48 CSVPSRLMIGCANLVESMCALSCLOSPSYRCL -> NTSV  
 FT SEQUENCE 542 AA: 62090 MW: 7CD40087EFAA5C8A CRC64;  
 SQ MEDONDESEPRKNTLM (IN ISOFORM B).  
 Query Match 87.5%; Score 210; DB 1; Length 542;  
 Best Local Similarity 86.7%; Pred. No. 8.8e-21;  
 Matches 39; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 ALPARGDHIENILSTYDAKSLCAELVCKEWRVTSIDGLMKR 45  
 DB 122 ALPEGLDHIENILSTYDAKSLCAELVCKEWRVTSIDGLMKR 166  
 RESULT 4  
 MT30\_YEAST STANDARD; PRT; 640 AA.  
 AC P39014.  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE MET30 protein.  
 GN MET30 OR YIL046W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180-1A;  
 RX MEDLINE=96069360; PubMed=8524217;  
 RA Thomas D., Kuras L., Barbey R., Cherest H., Blaiseau P.L.,  
 RA Surdin-Kerjan Y.;  
 RT "Met30p, a yeast transcriptional inhibitor that responds to S-  
 RT adenosylmethionine, is an essential protein with WD40 repeats.";  
 RL Mol. Cell. Biol. 15:6526-6534(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Hornell T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,  
 RA Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.;  
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS  
 CC GENES EXPRESSION.  
 CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT  
 CC PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z46861; CAAB6905.1; -;  
 DR EMBL: L26505; AAA96717.1; -;  
 DR SGD: S0001308; MET30.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PS50181; FBOX; 1.

```

DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Methionine biosynthesis;
KW Cysteine biosynthesis; Repeat; WD repeat.
FT DOMAIN 181 227 F-BOX.
FT REPEAT 300 328 WD 1.
FT REPEAT 340 368 WD 2.
FT REPEAT 380 408 WD 3.
FT REPEAT 419 449 WD 4.
FT REPEAT 461 499 WD 5.
FT REPEAT 509 538 WD 6.
FT REPEAT 550 578 WD 7.
FT REPEAT 607 635 WD 8 (POTENTIAL).
FT CONFLICT 61 61 M -> I (IN REF. 1).
SQ SEQUENCE 640 AA; 72835 MW; 5135D4HCA2E1EB97 CRC64;

Query Match 37.9%; Score 91; DB 1; Length 640;
Best Local Similarity 50.0%; Pred. No. 0.00014;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 14 ILSYLDKSLCAELVCKEYRVTSQGLMK 43
Db 195 ILSYLDKSLCAELVCKEYRVTSQGLMK 224

RESULT 5
POF1_SCHPO STANDARD; PRT; 605 AA.
ID POF1_SCHPO
AC P87053;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof1 (Skp1-binding protein 1).
GN POF1 OR SBP1 OR SPAC57A10.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Katayama S., Toda T.;
RT "Systematic genome-wide analysis of F-box protein-encoding genes in
RL fission yeast."
RN Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA Jones P., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson S., Quail M.A., Rabinowitsch E.,
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckart G., Aert R., Koben J., Gymnopoulos B.,
RA Meltjens I., Vastreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez J., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

```

```

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Snopcewski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -!- SUBUNIT: PART OF A SCF (SKP1-CULIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; AB032410; BAA84528.1; -
DR EMBL; Z94864; CAB08168.1; -
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPOTEINRPT.
DR ProDom; PD000018; WD40; 5.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS0181; FBOX; 1.
DR PROSITE; PS50082; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW DBL conjugation; Repeat; WD repeat.
FT DOMAIN 107 153 F-BOX.
FT REPEAT 271 299 WD 1.
FT REPEAT 311 339 WD 2.
FT REPEAT 350 379 WD 3.
FT REPEAT 390 420 WD 4.
FT REPEAT 432 460 WD 5.
FT REPEAT 472 500 WD 6.
FT REPEAT 510 538 WD 7.
SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 36.7%; Score 88; DB 1; Length 605;
Best Local Similarity 44.4%; Pred. No. 0.00034;
Matches 16; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 10 IAEIISYLDKSLCAELVCKEYRVTSQGLMK 45
Db 117 ISFRIISYLDKSLCAELVCKEYRVTSQGLMK 152

RESULT 6
POF1_SCHPO STANDARD; PRT; 662 AA.
ID POF1_SCHPO
AC Q9P7M4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof10 (Skp1-binding protein 2).
GN POF10 OR SBP2 OR SPBC1703.06.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC MEDLINE=21678905; PubMed=11820777;
RA Ikebe C., Komihama K.-I., Toda T., Nakayama K.-I.;
RT "Isolation and characterization of a novel F-box protein Pof10 in
fission yeast."

```

RL Biochem. Biophys. Res. Commun. 290:1399-1407(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RX MEDLINE-21848401; PubMed-11859360;  
 RA Mood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gholts S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 3 WD REPEATS (TRP-ASP DOMAINS).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AB061725; BAB55636.1; -;  
 DR EMBL: AL136536; CAB66450.1; -;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR003903; UIM.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 1.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS; 1; FALSE\_NEG.  
 DR PROSITE: PS50082; WD\_REPEATS; 2; 1.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Ubl conjugation; Repeat; WD repeat.  
 KM DOMAIN 28 74 F-BOX.  
 FT REPEAT 215 260 WD 1.  
 FT REPEAT 263 302 WD 2.  
 FT REPEAT 429 468 WD 3.  
 SO SEQUENCE 662 AA; 74205 MW; 44061C20A61F602C CRC64;  
 OY Query Match 34.0%; Score 81.5; DB 1; Length 662;  
 DB Best Local Similarity 30.8%; Pred. No. 0.0029;  
 DB Matches 16; Conservative 9; Mismatches 12; Indels 15; Gaps 1;

RESULT 7  
 POP1\_SCHPO  
 ID POP1\_SCHPO STANDARD; PRT; 775 AA.  
 AC P87060;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE WD-repeat protein pop1.  
 GN POP1 OR SPBC262.18.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RX MEDLINE-97347242; PubMed-9203581;  
 RA Komihama K., Toda T.;  
 RT "Fission yeast WD-repeat protein pop1 regulates genome ploidy through  
 RT ubiquitin proteasome-mediated degradation of the CDK inhibitor Rum1  
 RT and the S-phase initiator Cdc18.";  
 RL Genes Dev. 11:1548-1560(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RX MEDLINE-21848401; PubMed-11859360;  
 RA Mood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gholts S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).  
 RN [3]  
 RP SUBUNIT.  
 RC STRAIN-972;  
 RX MEDLINE-9914318; PubMed-990507;  
 RA Komihama K.-I., Ochotorena I., Toda T.;  
 RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-  
 RT complexes together with cullin-1 in fission yeast SCF (Skip-cullin-1-  
 RT F-box) ubiquitin ligase.";  
 RL Genes Cells 3:721-735(1998).  
 CC -1- FUNCTION: Involved in maintenance of ploidy through proteasome  
 CC dependent degradation of CDK inhibitor rum1 and S-phase initiator  
 CC cdc18. Functions as a recognition factor for rum1 and cdc18, which  
 CC are subsequently ubiquitinated and targeted to the 26S proteasome  
 CC for degradation.  
 CC -1- SUBUNIT: Homodimer and heterodimer with pop2. Binds to pool and  
 CC cdc18.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y08391; CAA69671.1;  
 DR EMBL: AL022103; CAA17898.1;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 3.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PS0181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 3.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 4.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Repeat; WD repeat.  
 FT DOMAIN 298 345 F-BOX.  
 FT REPEAT 444 472 WD 1.  
 FT REPEAT 484 538 WD 2.  
 FT REPEAT 575 603 WD 3.  
 FT REPEAT 615 645 WD 4.  
 FT REPEAT 657 687 WD 5.  
 SQ SEQUENCE 775 AA; 87816 MW; B06EDBA46553EEC1 CRC64;

Query Match 33.5%; Score 80.5; DB 1; Length 775;  
 Best Local Similarity 47.2%; Pred. No. 0.0046;  
 Matches 17; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 10 IAEINILSYLDKSLCAELVCKEYR-VTSDGMLWK 44  
 DB 308 ITNLVLTLDAPSLCAVSGVSHHWKLVSNDELWK 343

RESULT 8  
 POPR SCHPO STANDARD; PRT; 506 AA.  
 AC 009655; Q9P7V1;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein popl1.  
 GN POP1 OR SPAC29F6.01 OR SPAC30.05.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harrison C.L., Toda T.;  
 RT "Systematic genome-wide analysis of F-box protein-encoding genes in  
 RT fission yeast."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RC [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RC MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Squoroos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Woilkeart G., Aert R., Robben J., Glynnoprez B.,  
 RA Woltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambull R., Punnett B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Ielaire V., Mottier S.,  
 RA Gallibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovskii G.V., Ussery D., Barrell B.G., Nurse P.;  
 RA "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).

CC -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -I- SIMILARITY: CONTAINS 7 WD\_REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AB061694; BAB55543.1;  
 DR EMBL: AL136538; CAB66464.1;  
 DR EMBL: Z66525; CAA91423.1;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 3.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 8.  
 DR PROSITE: PS0181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 3.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Repeat; WD repeat.  
 FT DOMAIN 70 116 F-BOX.  
 FT REPEAT 219 236 WD 1.  
 FT REPEAT 259 298 WD 2.  
 FT REPEAT 301 338 WD 3.  
 FT REPEAT 345 386 WD 4.  
 FT REPEAT 388 426 WD 5.  
 FT REPEAT 427 464 WD 6.  
 FT REPEAT 468 505 WD 7.  
 SQ SEQUENCE 506 AA; 58257 MW; CEF34D4EFBEC2E10 CRC64;

Query Match 32.9%; Score 79; DB 1; Length 506;  
 Best Local Similarity 35.1%; Pred. No. 0.0047;  
 Matches 13; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 8 DHAENILSYLDKSLCAELVCKEYR-VTSDGMLWK 44  
 DB 78 EEVSLRFVSYLDIDLCCKIMSKRMLLEDPGIWK 114

RESULT 9  
 SCOB\_EMENT STANDARD; PRT; 678 AA.  
 ID SCOB\_EMENT  
 AC 000659;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sulfur metabolite repression control protein.  
 GN SCONB OR MAPBI.

OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 RN NCBI\_TaxId=5072;  
 RP SEQUENCE FROM N.A.  
 RA Natorf R.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE  
 CC REPRESSION.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON 2 FAMILY OF WD-REPEAT  
 CC PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U21220; AAC15905.1;  
 CC InterPro: IPR001810; F-box.  
 CC InterPro: IPR001680; WD40.  
 CC Pfam: PF00400; WD40; 7.  
 CC Pfam: PF00646; F-box; 1.  
 CC PRINTS: PR00320; GPROTEINRPT.  
 CC PRODOM: PD000018; WD40; 4.  
 CC SMART: SM00256; FBOX; 1.  
 CC SMART: SM00320; WD40; 7.  
 CC PROSITE: PS50181; FBOX; 1.  
 CC PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 CC PROSITE: PS0082; WD\_REPEATS\_2; 7.  
 CC PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 CC TRANSCRIPTION REGULATION Repeat; WD repeat.  
 CC DOMAIN 178 224 F-BOX.  
 CC FT REPEAT 347 375 WD 1.  
 CC FT REPEAT 387 415 WD 2.  
 CC FT REPEAT 427 455 WD 3.  
 CC FT REPEAT 466 496 WD 4.  
 CC FT REPEAT 508 543 WD 5.  
 CC FT REPEAT 553 595 WD 6.  
 CC FT REPEAT 607 635 WD 7.  
 CC FT REPEAT 647 675 WD 8.  
 CC SEQUENCE 678 AA; 76070 MW; D8400452E37B4C53 CRC64;  
 SQ  
 Query Match 32.9%; Score 79; DB 1; Length 678;  
 Best Local Similarity 40.0%; Pred. No. 0.0064;  
 Matches 18; Conservative 5; Mismatches 18; Indels 4; Gaps 1;  
 OY 1 ALPARGLDIAENILSYLDAKSLCAELVCKEWMYRTSDGMLMK 45  
 DB 183 ALP-----PEIAFKILCYLDPTSLCKASOVSRGMRALADDVWHR 223  
 RESULT 10  
 YKK7\_CABEL  
 ID YKK7\_CABEL STANDARD; PRT; 461 AA.  
 AC P34284;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical F-box/LRR-repeat protein C02F5.7 in chromosome III.  
 GN C02F5.7  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 RN NCBI\_TaxId=6239;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;

RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,  
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaughan K.,  
 RA Waterston J., Watson A., Weinstock L., Wilkinson-Spratt J.,  
 RA Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L14745; AAA27922.2;  
 CC PIR: S44609; S44609.  
 CC MOPREP: C02F5.7; CE24780.  
 CC InterPro: IPR001810; F-box.  
 CC InterPro: IPR003592; LRR\_out.  
 CC Pfam: PF00646; F-box; 2.  
 CC SMART: SM00256; FBOX; 1.  
 CC SMART: SM00370; LRR; 5.  
 CC PROSITE: PS50181; FBOX; 1.  
 CC PROSITE: PS50181; FBOX; 1.  
 CC HYPOTHETICAL protein; Repeat; Leucine-rich repeat.  
 CC DOMAIN 54 100 F-BOX.  
 CC FT REPEAT 122 147 LRR 1.  
 CC FT REPEAT 148 173 LRR 2.  
 CC FT REPEAT 174 199 LRR 3.  
 CC FT REPEAT 200 225 LRR 4.  
 CC FT REPEAT 226 251 LRR 5.  
 CC FT REPEAT 252 277 LRR 6.  
 CC FT REPEAT 278 303 LRR 7.  
 CC FT REPEAT 304 329 LRR 8.  
 CC FT REPEAT 330 355 LRR 9.  
 CC FT REPEAT 356 381 LRR 10.  
 CC FT REPEAT 408 433 LRR 11.  
 CC SEQUENCE 461 AA; 51552 MW; BB28C98A5CA7B1D5 CRC64;  
 SQ  
 Query Match 32.5%; Score 78; DB 1; Length 461;  
 Best Local Similarity 37.5%; Pred. No. 0.0058;  
 Matches 12; Conservative 8; Mismatches 12; Indels 0; Gaps 0;  
 OY 14 ILSYDAKSLCAELVCKEWMYRTSDGMLMK 45  
 DB 68 VESFLDTKALCRSAQVCRSMILADGSMOR 99  
 RESULT 11  
 CC4\_CANAL  
 ID CC4\_CANAL STANDARD; PRT; 684 AA.  
 AC P53699;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cell division control protein 4.  
 GN CDC4.  
 OS Candida albicans (Yeast).

CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SGY126;  
 RA Shieh J.C., White A.M., Rosamond J.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA  
 CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE  
 CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD  
 CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND  
 CC VARIOUS ASPECTS OF SPOGULATION. REQUIRED FOR HTAI-HTBI LOCUS  
 CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X96763; CAA65538.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPOTETINBRPT.  
 DR ProDom: PD000018; WD40; 1.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 4.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Cell division; Mitosis; Sporulation; Repeat; WD repeat.  
 KW DOMAIN 212 258  
 FT REPEAT 322 351 WD 1.  
 FT REPEAT 363 391 WD 2.  
 FT REPEAT 403 431 WD 3.  
 FT REPEAT 442 468 WD 4 (POTENTIAL).  
 FT REPEAT 478 506 WD 5.  
 FT REPEAT 519 549 WD 6.  
 FT REPEAT 561 589 WD 7.  
 SQ SEQUENCE 684 AA: 76090 MW: 3DD65DB31293B107 CRC64;  
 Query Match 32.1%; Score 77; DB 1; Length 684;  
 Best Local Similarity 38.9%; Pred. No. 0.012;  
 Matches 14; Conservative 8; Mismatches 14; Indels 0; Gaps 0;  
 OY 10 IAEILSYLDKSLCAELVCKEYRVTSIGMLMK 45  
 Db 222 VTKMLSYLDYKTLSSVAQVCKRMEDINPDWIK 257  
 RESULT 12  
 FBX8\_HUMAN  
 ID FBX8\_HUMAN STANDARD; PRT; 319 AA.  
 AC Q9NRD0; Q9NRD5; Q9URC4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box only protein 8 (F-box/SEC7 protein FBS) (DC10).  
 GN FBX08 OR FBX8 OR FBS.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.

RX MEDLINE-20003060; PubMed=10531035;  
 RA Cenciarelli C., Chiar D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-2039565; PubMed=10945468;  
 RA Llyin G.P., Riialand M., Pigeon C., Guen Gulllouz C.;  
 RT "cDNA cloning and expression analysis of new members of the mammalian  
 RT F-box protein family.";  
 RL Genomics 67:40-47(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Li Y., Peng Y., Qian B., Zhang Z., Han Z., Fu G., Chen Z.;  
 RT "Novel genes expressed in human dentritic cell.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY PROMOTE GUANINE-NUCLEOTIDE EXCHANGE ON AN ARF.  
 CC PROMOTES THE ACTIVATION OF ARF THROUGH REPLACEMENT OF GDP WITH  
 CC GTP (POTENTIAL).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SEC7 DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF174596; AAF04517.1; ALT\_INIT.  
 DR EMBL: AF233224; AAF67154.1; -  
 DR EMBL: AF201932; AAF86868.1; -  
 DR EMBL: BC014679; AAH14679.1; -  
 DR HSSP: Q99418; IPBV.  
 DR GeneW: HGNC:13587; FBX08.  
 DR MIM: 605649; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR000904; Sec7.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF01369; Sec7; 1.  
 DR SMART: SM00222; Sec7; 1.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS50190; SEC7; 1.  
 KW Guanine-nucleotide releasing factor.  
 FT DOMAIN 68 111  
 FT DOMAIN 146 276 SEC7.  
 FT CONFLICT 119 119 C->W (IN REF. 3).  
 FT CONFLICT 135 135 L->V (IN REF. 3).  
 SQ SEQUENCE 319 AA: 37068 MW: 125D6DB57CA7C79 CRC64;  
 Query Match 27.7%; Score 66.5; DB 1; Length 319;  
 Best Local Similarity 45.2%; Pred. No. 0.14;  
 Matches 14; Conservative 5; Mismatches 9; Indels 3; Gaps 1;  
 OY 14 ILSYLDKSLCAELVCKEYRVTSIGMLMK 44  
 Db 82 ILSYLDKSLCAELVCKEYRVTSIGMLMK 109  
 RESULT 13  
 FBX8\_MOUSE  
 ID FBX8\_MOUSE STANDARD; PRT; 319 AA.  
 AC Q90ZNS;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20399565; PubMed=10945468;
RA Ilyin G.P., Riialand M., Pigeon C., Guquen-Guillouzo C.;
RT "cDNA cloning and expression analysis of new members of the mammalian
RL F-box protein family.";
RN Genomics 67:40-47(2005).
RP [3]
RC SEQUENCE OF 121-412 FROM N.A.
RD TISSUE=Pancreas;
RE Strausberg R.;
RF Submitted (MAY-2001) to the EMBL/GenBank/DDIJ databases.
RG -I- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHORYLATED
RH PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION. LIKELY
RI TO BE INVOLVED IN KEY SIGNALING PATHWAYS CRUCIAL FOR NORMAL LIMB
RJ DEVELOPMENT.
RK -I- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
RL (BY SIMILARITY).
RM -I- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, KIDNEY, LUNG AND LIVER.
RN -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
RO -I- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
RR -----
RS This SWISS-PROT entry is copyright. It is produced through a collaboration
RW between the Swiss Institute of Bioinformatics and the EMBL Outstation *
RX the European Bioinformatics Institute. There are no restrictions on its
RY use by non-profit institutions as long as its content is in no way
RZ modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See http://www.isb-sib.ch/announce/
RB or send an email to license@isb-sib.ch).
RC -----
RD EMBL: AF281859; AACG22739.1; -
RE EMBL: BC007380; AAH07380.1; -
RF Genew: HGNC:10847; SHFW3.
RG MIM: 6000095; -
RH InterPro: IPR001810; F-box.
RI InterPro: IPR001680; WD40.
RJ Pfam: PF004400; WD40_4.
RK Pfam: PF00646; F-box; 1.
RL SMART: SM00320; WD40_5.
RM SMART: SM00256; FBox; 1.
RN PROSITE: PS0181; FBox; 1.
RO PROSITE: PS00678; WD_REPEATS_1; FALSE_NEG.
RW PROSITE: PS50082; WD_REPEATS_2; 1.
RX PROSITE: PS50294; WD_REPEATS_REGION; 1.
RY Developmental protein; Ub1 conjugation pathway; Repeat; WD repeat.
RZ DOMAIN 25 71 F-Box.
RA FT REPEAT 161 198 WD 1.
RB FT REPEAT 200 237 WD 2.
RC FT REPEAT 291 329 WD 3.
RD FT REPEAT 335 374 WD 4.
RE FT REPEAT 363 363 L -> P (IN REF. 3).
RF SEQUENCE 412 AA; 46337 MW; 36C175FD830D83B CRC64;
RG
RH Query Match 26.7%; Score 64; DB 1; Length 412;
RI Best Local Similarity 37.3%; Pred. NO. 0.4;
RJ Matches 19; Conservative 8; Mismatches 18; Indels 6; Gaps 2.
RK
RL QY 1 ALPARG--LDHTAENLL-----SYLAKSLCAAEIVCKEYRVTSQMLMKK 45
RN | || | | | : | : | | | | : | : | | | : | : | : | :
RU Db 20 ARPAGPALWRLPBELLLITCSYLDMRALGRLAOVCRWLRRFTSCDLIMKR 70
RV
RW RESULT 15
RX SCQ2_NEUCR
RY ID SCQ2_NEUCR STANDARD: PRT: 650 AA.
RZ AC Q01277;
RA DT 15-JUL-1998 (Rel. 36, Created)
RB DT 15-JUL-1998 (Rel. 36, Last sequence update)
RC DT 16-OCT-2001 (Rel. 40, Last annotation update)
RD DE Sulfur controller-2 (SCON2).
RE GN SCON-2.
RF OS Neurospora crassa.
RG OC Eukaryota; Fungi; Ascomycota; Pezilomyces; Sordariomycetes;

```

OC Sordariales; Sordariaceae; Neurospora.  
 OX NCHL\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A;  
 RX MEDLINE=95241499; PubMed=7724564;  
 RA Kumar A., Paletta J.V.;  
 RT "The sulfur controller-2 negative regulatory gene of Neurospora  
 crassa encodes a protein with beta-transducin repeats.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).  
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.  
 CC -1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONE/SCON-2 FAMILY OF WD-REPEAT  
 CC PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U17251; AAA68968.1; -.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR Pfam; PF00646; F-box; 1.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR PRODOM; PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS00181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Transcription regulation; Repeat; WD repeat.  
 FT DOMAIN 124 170  
 FT REPEAT 292 320 WD 1.  
 FT REPEAT 332 360 WD 2.  
 FT REPEAT 372 400 WD 3.  
 FT REPEAT 411 441 WD 4.  
 FT REPEAT 453 488 WD 5.  
 FT REPEAT 528 564 WD 6.  
 FT REPEAT 576 604 WD 7.  
 FT REPEAT 616 644 WD 8.  
 SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;  
 Query Match 25.8%; Score 62; DB 1; Length 650;  
 Best Local Similarity 32.6%; Pred. No. 1.2;  
 Matches 14; Conservative 6; Mismatches 19; Indels 4; Gaps 1;  
 OY 1 ALPARGIDHIAENILSYDAKSLCAAEIVCKEYRYVTSQGLM 43  
 Db 129 ALPV---ELAQKVLCLDTVSLTKAAQVSGRRRTLADSDAVM 167

Search completed: April 10, 2003, 13:19:32  
 Job time : 4.23386 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 : Search time 10.6334 Seconds  
(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_147\_191  
Perfect score: 240  
Sequence: 1 ALPARGLDHIAENILSYLDA.....AELVCKEYRWISDGMWK 45

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL.21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_todent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	100.0	569	11	Q92159 mus musculu
2	240	100.0	569	11	Q90015 mus musculu
3	236	98.3	569	11	Q91197 mus musculu
4	216	90.0	62	13	Q800N7 xenopus lae
5	210	87.5	563	11	Q923H0 mus musculu
6	188	78.3	510	5	Q9VDE3 mus musculu
7	184	76.7	510	5	Q44382 drosophila
8	83.5	34.8	619	10	Q91TX2 arabidopsis
9	83	34.6	483	4	Q94926 homo sapien
10	83	34.6	491	4	Q90UT9 homo sapien
11	82	34.2	464	5	Q9V605 drosophila
12	80	33.3	144	11	Q9Q2N2 mus musculu
13	78	32.5	197	4	Q9UKC2 homo sapien
14	78	32.5	466	5	Q8T3G0 caenorhabdi
15	78	32.5	511	5	Q8TOK1 drosophila
16	77.5	32.3	635	10	Q9A0H6 populus tre

17	77	32.1	474	4	Q9UK66 homo sapien
18	77	32.1	496	4	Q9Y2K7 homo sapien
19	77	32.1	535	4	Q9UKC8 homo sapien
20	77	32.1	636	4	Q9XN22 homo sapien
21	77	32.1	674	4	Q9P010 homo sapien
22	77	32.1	690	4	Q90UT7 homo sapien
23	77	32.1	691	4	Q9P0X5 homo sapien
24	77	32.1	694	4	Q9UKA1 homo sapien
25	77	32.1	757	4	Q9H7H5 homo sapien
26	77	32.1	1182	5	Q9VH60 drosophila
27	77	32.1	1326	5	Q9VZF4 drosophila
28	75.5	31.5	614	10	Q9STV5 arabidopsis
29	75.5	31.5	623	10	Q8RM08 arabidopsis
30	75	31.2	157	10	Q9LUB6 caenorhabdi
31	75	31.2	518	5	Q968Y8 caenorhabdi
32	75	31.2	535	5	Q968Y7 caenorhabdi
33	75	31.2	535	5	Q968Y7 caenorhabdi
34	74	30.8	151	10	Q9STX3 arabidopsis
35	74	30.8	192	4	Q90HT1 homo sapien
36	74	30.8	253	13	Q91B68 brachydanio
37	74	30.8	392	13	Q91BH3 brachydanio
38	74	30.4	508	5	Q9VAK7 drosophila
39	72.5	30.2	410	4	Q9H6V9 homo sapien
40	72.5	30.2	424	4	Q8TDB1 homo sapien
41	72.5	30.2	435	4	Q13309 homo sapien
42	72.5	30.2	442	4	Q8TDB2 homo sapien
43	72.5	30.0	135	10	Q9LU04 arabidopsis
44	71	29.6	573	11	Q9QXW2 mus musculu
45	70.5	29.4	592	10	Q04197 arabidopsis

## ALIGNMENTS

RESULT 1  
ID Q92159 PRELIMINARY: PRT: 569 AA.  
AC Q92159;  
DC 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DR 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Beta-transducin repeat containing protein.  
GN BTRC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. PubMed=990853;  
RA MEDLINE=99145465; PubMed=990853;  
RA Spencer E., Jiang J., Chen Z.,  
RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein  
RT Slnb/beta-Trcp.";  
RL Genes Dev. 13:284-294(1999).  
CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
DR EMBL; AF112979; AAD04181.1;  
DR MGD; MGI:1338871; BTRC.  
DR InterPro; IPR001810; F-box.  
DR InterPro; IPR001800; WD40.  
DR Pfam; PF00646; F-box; 1.  
DR Pfam; PF00400; WD40; 7.  
DR PRINTS; PR00320; GPROTEINRPT.  
DR Prodom; PD000018; WD40; 4.  
DR SMART; SM00256; FBOX; 1.  
DR PROSITE; PS00320; WD40; 7.  
DR PROSITE; PS00181; FBOX; 1.  
DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
DR PROSITE; PS00082; WD\_REPEATS\_2; 7.  
DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 569 AA: 65047 MW: BC7C7A44815BD96 CRC64:  
Query Match 100.0%; Score 240; DB 11; Length 569;

Best Local Similarity 100.0%; Pred. No. 3,1e-25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 45  
DB 147 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 191

RESULT 2  
Q90U15 PRELIMINARY; PRT; 569 AA.  
ID Q90U15  
AC Q90U15  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)  
DE ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)  
DE (F-box WD40 repeat protein 1).  
GN BTRC OR FBXW1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCB1\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99199275; PubMed=10097128;  
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,  
RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,  
RA Nakayama K.-I.,  
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a  
RT ubiquitin ligase SKP1/Cul 1/F-box protein FWD1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99075339; PubMed=9859996;  
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,  
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,  
RT "Identification of the receptor component of the IkappaBalpha-  
RT ubiquitin ligase.";  
RL Nature 396:590-594(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.,  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RX MEDLINE=21601157; PubMed=11735228;  
RA Maryama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,  
RA Nakayama K.-I.,  
RT "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of  
RT Caenorhabditis elegans SEL-10.";  
RL Genomics 78:214-222(2001).  
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
DR EMBL: AF081887; AAD1775.1; -;  
DR EMBL: AF099932; AAD08701.1; -;  
DR EMBL: BC003989; AAH03989.1; -;  
DR EMBL: AF391190; AAL40929.1; -;  
DR EMBL: AF391178; AAL40929.1; JOINED.  
DR EMBL: AF391179; AAL40929.1; JOINED.  
DR EMBL: AF391180; AAL40929.1; JOINED.  
DR EMBL: AF391181; AAL40929.1; JOINED.  
DR EMBL: AF391182; AAL40929.1; JOINED.  
DR EMBL: AF391183; AAL40929.1; JOINED.  
DR EMBL: AF391184; AAL40929.1; JOINED.  
DR EMBL: AF391185; AAL40929.1; JOINED.  
DR EMBL: AF391186; AAL40929.1; JOINED.  
DR EMBL: AF391187; AAL40929.1; JOINED.  
DR EMBL: AF391188; AAL40929.1; JOINED.  
DR EMBL: AF391189; AAL40929.1; JOINED.  
DR MGD: MGI:1338871; Btrc.  
DR InterPro: IPR001810; F-box  
DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 7.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR ProDom: PD000018; WD40; 4.  
DR SMART: SM00256; FBOX; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS50181; FBOX; 1.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW ligase; Repeat; WD repeat.  
SQ SEQUENCE 569 AA; 65209 MW; 6C7D6544815B2296 CRC64;

Query Match 100.0%; Score 240; DB 11; Length 569;  
Best Local Similarity 100.0%; Pred. No. 3,1e-25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 45  
DB 147 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 191

RESULT 3  
Q9R1G7 PRELIMINARY; PRT; 569 AA.  
ID Q9R1G7  
AC Q9R1G7  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)  
DE Beta-transducin repeat-containing protein.  
GN BTRC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCB1\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Winston J., Ellledge S.J., Harper J.W.,  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
DR EMBL: AF110396; AAD41025.1; -;  
DR MGD: MGI:1338871; Btrc.  
DR InterPro: IPR001810; F-box.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00646; F-box; 1.  
DR Pfam: PF00400; WD40; 7.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR ProDom: PD000018; WD40; 4.  
DR SMART: SM00256; FBOX; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS50181; FBOX; 1.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

Query Match 98.3%; Score 236; DB 11; Length 569;  
Best Local Similarity 97.8%; Pred. No. 1,1e-24;  
Matches 44; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 45  
DB 147 ALPARGLDHIAENILSYLDAKSLCAAEVCKEYRVTSIDGMLMK 191

RESULT 4  
Q8UUN7 PRELIMINARY; PRT; 62 AA.  
ID Q8UUN7  
AC Q8UUN7  
DT 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)  
DE Beta-Trip protein (Fragment).

GN BETA-TRCP.  
 OS Xenopus laevis (African clawed frog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
 CC Xenopodinae; Xenopus.  
 ON NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Carnevall F.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ballarino M.;  
 RL "Analisi strutturale e funzionale del gene beta-trcp in Xenopus laevis";  
 RL Thesis (2001),  
 RL Department of Genetics and Molecular Biology Charles Darwin,  
 RL University of Rome la Sapienza, Rome, Italy.  
 DR EMBL: AJ428934; CAD21911.1; -  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF00646; F-box; 1.  
 DR SMART: SM00256; FBOX; 1.  
 DR PROSITE: PS50181; FBOX; 1.  
 FT NON\_TER  
 FT 1 62  
 FT 62 62  
 SEQUENCE 62 AA: 7299 MW: A735E62CA4A2C02 CRC64;

Query Match 90.0%; Score 216; DB 13; Length 62;  
 Best Local Similarity 95.2%; Pred. No. 5.4e-23;  
 Matches 40; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 ARGLDHIAENITSLYDAKSLCAELVCKEMRYTSDGMLMK 45  
 DB 1 ARGLDHIAENITSLYDAKSLCAELVCKEMRYTSDGMLMK 42

RESULT 5  
 O923H0 PRELIMINARY; PRT; 563 AA.  
 AC 0923H0;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE F-box/WD40 repeat-containing protein HOS.  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shraim-PVB/N;  
 RL "Mouse homolog of HOS (mHOS) is overexpressed in skin tumors and implicated in constitutive activation of NF-kappaB";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: A038079; AAK72095.1; -  
 DR MGD: MG1:2144023; Fbxw1b.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001860; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR ProDom: PD000018; WD40; 4.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Repeat; WD repeat.  
 SQ SEQUENCE 563 AA: 64741 MW: 9AB5625F3F5E3496 CRC64;

Query Match 87.5%; Score 210; DB 11; Length 563;  
 Best Local Similarity 86.7%; Pred. No. 5.2e-21;

Matches 39; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ALPARGLDHIAENITSLYDAKSLCAELVCKEMRYTSDGMLMK 45  
 DB 143 ALPARGLDHIAENITSLYDAKSLCAELVCKEMRYTSDGMLMK 187

RESULT 6  
 O9VDE3 PRELIMINARY; PRT; 510 AA.  
 ID 09VDE3;  
 AC 09VDE3;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE SLMB protein (SLMB).  
 GN SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-BERKELEY;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Chame M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-U., Andrews-Plamkoc C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dev I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jaleel M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Jimalal B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,  
 RA Palazzo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Sytkus R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao O.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).

[12] SEQUENCE FROM N.A.  
 RP TISSUE-OVARY, AND IMAGINAL DISC;  
 RC MEDLINE=20245299; PubMed=10781936;  
 RX Miletich I., Limbourg-Bouchon B.;  
 RA "Drosophila null slmb clones transiently deregulate Hedgehog-independent transcription of wingless in all limb discs, and induce decapentaplegic transcription linked to imaginal disc regeneration";  
 RT Mech. Dev. 93:15-26(2000).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

DR EMBL: AE003733; AAF55853.1; -  
 DR EMBL: AF222924; AAF63214.1; -  
 DR EMBL: AF222923; AAF63213.1; -  
 DR FlyBase: FBgn0023423; slmb.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; F-box; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; F-box; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 510 AA; 58952 MW; F4D5DF126F58A012 CRC64;

Query Match 78.3%; Score 188; DB 5; Length 510;  
 Best Local Similarity 79.5%; Pred. No. 5.9e-18;  
 Matches 35; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 LPARGDHIENILSYDANSICAEELVCKEYRVTSQGMIMKK 45  
 Db 92 LPKIGLDHIGENILSYDAESLKSSELVCKEYRVTSQGMIMKK 135

RESULT 7  
 ID 044382 PRELIMINARY; PRT; 510 AA.  
 AC 044382;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE SLMB.  
 GN SLMB OR SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98121115; PubMed=9461217;  
 RA Jiang J., Struhl G.;  
 RT Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slmb.  
 RL Nature 391:493-496(1998).  
 CC -1- SIMILARITY: CONTAINS 7 WD\_REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF032878; AAC38852.1; -  
 DR FlyBase: FBgn0023423; slmb.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; F-box; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 510 AA; 58952 MW; DB0243D3730A5E8 CRC64;

Query Match 76.7%; Score 184; DB 5; Length 510;  
 Best Local Similarity 77.3%; Pred. No. 2.2e-17;  
 Matches 34; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 2 LPARGDHIENILSYDANSICAEELVCKEYRVTSQGMIMKK 45  
 Db 92 LPKIGLDHIGENILSYDAESLKSSELVCKEYRVTSQGMIMKK 135

RESULT 8  
 ID 09LTX2 PRELIMINARY; PRT; 619 AA.  
 AC 09LTX2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Transport inhibitor response 1 protein (At5g49980/K9P8\_12).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneo T., Katoh T., Asamizu E., Kotani H.,  
 RA Tabata S.;  
 RT Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty p1 and TAC clones.  
 RT DNA Res. 7:31-63(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Baeh J.,  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narisaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.D., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Exner J.R.;  
 RT "Arabidopsis cDNA clones."  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB024032; BAA97019.1; -  
 DR EMBL: AY056431; AAU08287.1; -  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF00646; F-box; 1.  
 SQ SEQUENCE 619 AA; 69316 MW; D0614AF071EB4PD2 CRC64;

Query Match 34.8%; Score 83.5; DB 10; Length 619;  
 Best Local Similarity 45.9%; Pred. No. 0.0043;  
 Matches 17; Conservative 8; Mismatches 7; Indels 5; Gaps 2;

OY 3 PARIGDHIENILSYDANSICAEELVCKEYRVTSQGMIMKK 36  
 Db 51 PDHVLNVLNVLQFLDSR--CDRNAASLVCKSMWVRV 85

RESULT 9  
 ID 094926 PRELIMINARY; PRT; 483 AA.  
 AC 094926;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE KIAA0840 protein (Fragment).  
 GN KIAA0840 OR FBL7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=99156230; PubMed=10048485;  
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirosewa M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code

```

RT for large proteins in vitro."
RL DNA Res. 5:355-364(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20003060; PubMed=10531035;
RA Cenciarelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,
RT Pagano M.;
RT "Identification of a family of human F-box proteins."
RL Curr. Biol. 9:1177-1179(1999).
DR EMBL: AB020647; BAA74863.1; -
DR EMBL: AF174593; AAF04514.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR003592; LRR_out.
DR Pfam: PF00646; F-box; 1.
DR SMART: SM00256; FBOX; 1.
DR PROSITE: PSS0181; FBOX; 1.
DR NONTER
SQ SEQUENCE 483 AA; 53773 MW; AEC3991439E396EF CRC64;

Query Match 34.6%; Score 83; DB 4; Length 483;
Best Local Similarity 37.8%; Pred. No. 0.0037;
Matches 14; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 8 DHIAENILSYLDAKSLCAELVCKEYRYVTSQGLMK 44
DB 111 DHSWQIFSLPTNOLCRCAVRCRRWYNLAMPRLMR 147

RESULT 10
Q9UT9 PRELIMINARY; PRT; 491 AA.
AC Q9UT9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F-box protein FBL6.
CN FBL6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20399565; PubMed=10945468;
RA Llyin G.P., Rialland M., Pigeon C., Gugen-Guillouzo C.;
RT "cDNA Cloning and Expression Analysis of New Members of the Mammalian
RT F-box Protein Family."
RL Genomics 67:40-47(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Llyin G.P.;
RT "F-box protein FBL6 containing leucine-rich repeats."
RT Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF199356; AAF09248.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR003592; LRR_out.
DR Pfam: PF00646; F-box; 1.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00370; LRR; 1.
DR PROSITE: PSS0181; FBOX; 1.
SQ SEQUENCE 491 AA; 54574 MW; A70C092A4E748C69 CRC64;

Query Match 34.6%; Score 83; DB 4; Length 491;
Best Local Similarity 37.8%; Pred. No. 0.0038;
Matches 14; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 8 DHIAENILSYLDAKSLCAELVCKEYRYVTSQGLMK 44
DB 119 DHSWQIFSLPTNOLCRCAVRCRRWYNLAMPRLMR 155

RESULT 11

```

```

Q9V605
ID Q9V605 PRELIMINARY; PRT; 464 AA.
AC Q9V605;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CG9003 protein.
CN CG9003
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Beeson K.Y., Bonos P.V., Borker J., Brokstein P., Brotlier P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brokstein P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Kadir C., Kravitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Styksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-T., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AB003625; AAF58635.1; -
DR FlyBase: FBgn0033639; CG9003.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00560; LRR; 2.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00370; LRR; 9.
DR PROSITE: PSS0181; FBOX; 1.
SQ SEQUENCE 464 AA; 51606 MW; 18B0C987BCD20841 CRC64;

Query Match 34.2%; Score 82; DB 5; Length 464;
Best Local Similarity 46.9%; Pred. No. 0.0049;
Matches 15; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

QY 14 ILSYLDAKSLCAELVCKEYRYVTSQGLMK 45

```

Db 64 VESFLDVSICRCQAQVCKYNNVIALDSSWOK 95

RESULT 12

ID 090ZN2 PRELIMINARY; PRT; 144 AA.

AC 090ZN2;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

GN F-box protein FBX12 (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Winston J.T., Elledge S.J., Harper W.;

RT "A family of Mammalian F-box Proteins."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF176528; AAF09137.1; -

DR MGD: MGI:1354703; FBX012.

DR InterPro: IPR001810; F-box.

DR Pfam: PF00646; F-box; 1.

DR SMART: SM00256; FBOX; 1.

DR PROSITE: PS50181; FBOX; 1.

FT NON\_TER 1

SEQUENCE 144 AA; 17265 MW; C02B7CC36348AEDA CRC64;

Query Match 33.3%; Score 80; DB 11; Length 144;

Best Local Similarity 53.1%; Pred. No. 0.0024;

Matches 17; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 14 ILSTLDKSLCAELVCKEYRVTSDGMLMK 45

Db 17 ILSTLDKSLCAELVCKEYRVTSDGMLMK 48

RESULT 13

ID 090KC2 PRELIMINARY; PRT; 197 AA.

AC 090KC2;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE F-box protein FBX11 (Fragment).

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-20003060; PubMed-10531035;

RA Cenciarelli C., Chiur D.S., Guardavaccaro D., Parks W., Vidal M.,

RA Pagano M.;

RT "Identification of a family of human F-box proteins."

RL Curr. Biol. 9:1177-1179(1999).

DR EMBL: AF174599; AAF04520.1; -

DR InterPro: IPR001810; F-box.

DR InterPro: IPR003267; Pro-rich.

DR Pfam: PF00646; F-box; 1.

DR PRINTS: PR01217; PRICHTEXTENS.

DR PRINTS: PR00021; PROTRICH.

DR SMART: SM00256; FBOX; 1.

DR PROSITE: PS50181; FBOX; 1.

FT NON\_TER 1

SEQUENCE 197 AA; 22336 MW; F8C37A23510924D0 CRC64;

Query Match 32.5%; Score 78; DB 4; Length 197;

Best Local Similarity 39.5%; Pred. No. 0.0066;

Matches 15; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 8 DHIAENILSTLDKSLCAELVCKEYRVTSDGMLMK 45

Db 146 DEVVLKIFSYLLEDDICRACVCKRSESLANDPMLMK 183

RESULT 14

ID 08T3G0 PRELIMINARY; PRT; 466 AA.

AC 08T3G0;

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DE Hypothetical 52.1 kDa protein.

GN C02F5.7.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-BRISTOL N2;

RC MEDLINE=99069613; PubMed=9851916;

DR Waterston R.;

RT "Genome sequence of the nematode C. elegans: a platform for

RT investigating biology. The C. elegans Sequencing Consortium."

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA STRAIN-BRISTOL N2;

RC Anderson K.;

RT "The sequence of C. elegans cosmid C02F5.7."

RL Submitted (May-1993) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA STRAIN-BRISTOL N2;

RC Waterston R.;

RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: L14745; AAM15540.1; -

DR Hypothetical protein.

SEQUENCE 466 AA; 52064 MW; 05BCEB2AA606E15F CRC64;

Query Match 32.5%; Score 78; DB 5; Length 466;

Best Local Similarity 37.5%; Pred. No. 0.018;

Matches 12; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 14 ILSTLDKSLCAELVCKEYRVTSDGMLMK 45

Db 68 VESFLDTRKALCRSAQVCKRSMIALDSSNMQR 99

RESULT 15

ID 08TOK1 PRELIMINARY; PRT; 511 AA.

AC 08TOK1;

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE G126184P.

GN C02010.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

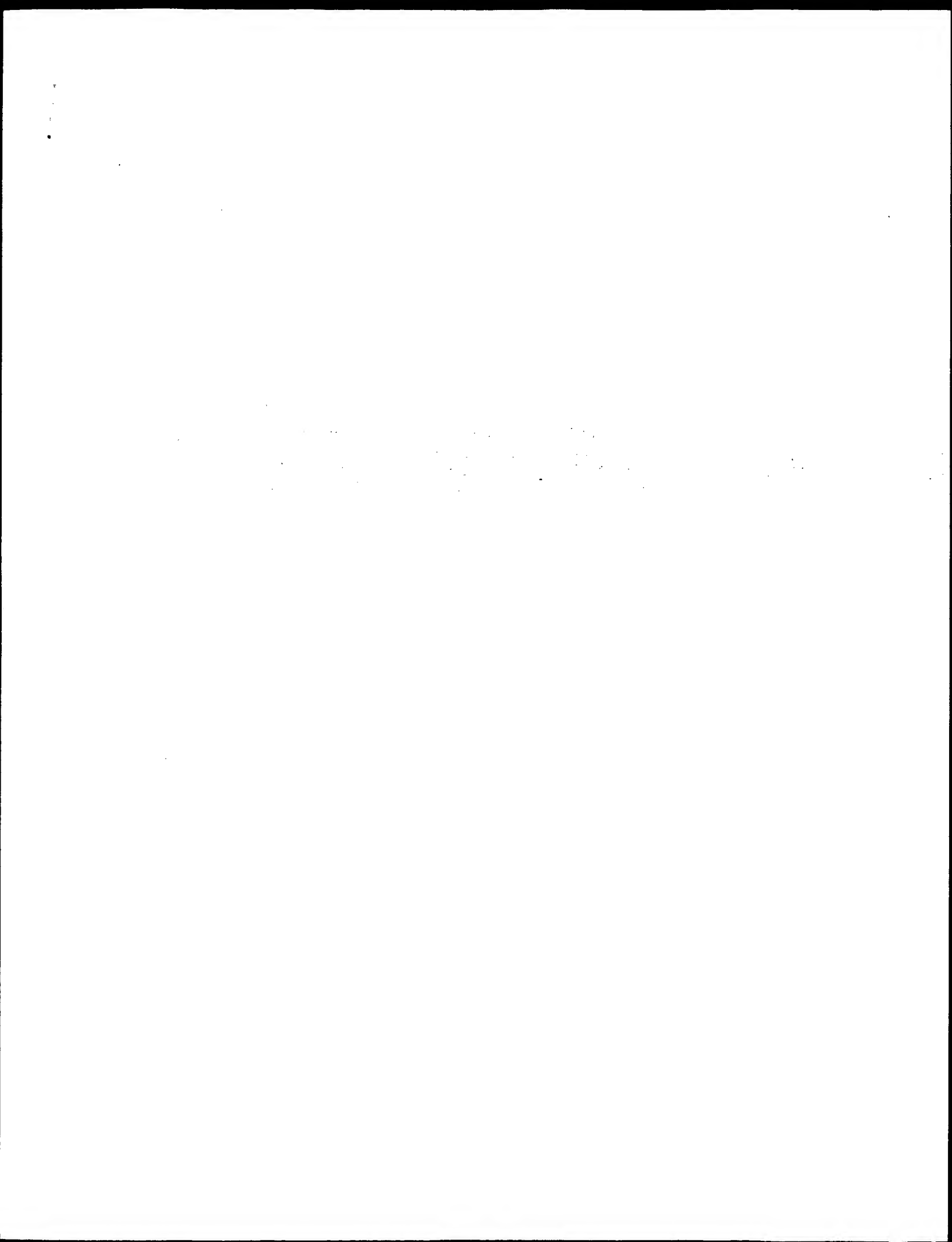
RA STRAIN-BERKELEY;

RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,

RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,







GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:20:44 ; Search time 10.4775 Seconds  
(without alignments)  
432.406 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_259\_292  
Perfect score: 186  
Sequence: 1 RHCRSETSGKGYCLQYDQKIVSGLRNTIKIW 34

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

A.GeneSeq\_101002:\*

- 1: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*
- 2: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*
- 3: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*
- 4: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*
- 5: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*
- 6: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*
- 7: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*
- 8: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*
- 9: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*
- 10: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*
- 11: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*
- 12: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*
- 13: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*
- 14: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*
- 15: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*
- 16: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*
- 17: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*
- 18: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*
- 19: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*
- 20: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*
- 21: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*
- 22: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*
- 23: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	517	16 AAR85852	WD-40 domain-contg
2	186	100.0	569	20 AAY24054	A human beta-trans
3	186	100.0	569	21 AAB12812	Mouse ubiquitin li
4	186	100.0	569	21 AAB12813	Human beta-transdu
5	186	100.0	569	21 AAY96697	Human beta-TRCP
6	186	100.0	569	21 AAY83041	F-box protein FBP-
7	186	100.0	569	21 AAY83250	F-box protein FBP-
8	186	100.0	569	21 AAY83250	F-box protein FBP-
9	186	100.0	569	21 AAY42429	Human cell signal
10	186	100.0	569	22 AAB48298	Human ZF11 protein

11	186	100.0	579	22 AAM78583	Human protein SEQ
12	186	100.0	590	22 AAM00847	Human bone marrow
13	186	100.0	605	22 AAM78582	Human protein SEQ
14	186	100.0	608	22 AAM00960	Human bone marrow
15	186	100.0	632	22 AAM78584	Human protein SEQ
16	186	100.0	654	22 AAM79566	Human protein SEQ
17	186	100.0	654	22 AAM79567	Human protein SEQ
18	186	100.0	654	22 AAM79568	Human protein SEQ
19	181	97.3	34	16 AAR84935	Peptide ri from a
20	167	89.8	510	22 AAB59857	Drosophila melanog
21	165	88.7	542	21 AAY96696	Human E3 ubiquitin
22	165	88.7	542	22 AAM79127	Human polypeptide
23	165	88.7	542	22 AAM40208	Human polypeptide
24	165	88.7	550	22 AAM41994	Human DNA repair a
25	123	66.1	265	22 AAB69442	Human protein SEQ
26	96	51.6	138	22 AAM80111	Peptide ri from a
27	87	46.8	29	16 AAR84939	Drosophila melanog
28	75	40.3	1326	22 AAB67237	Drosophila melanog
29	75	40.3	1326	22 AAB67238	Drosophila melanog
30	75	40.3	1326	22 AAB70051	F-box protein Met3
31	73	39.2	640	21 AAY83252	Human polypeptide
32	72	38.7	357	23 AAB90392	Human polypeptide
33	72	38.7	540	20 AAY22465	Human mammary sel-
34	72	38.7	540	20 AAY22468	Human mammary sel-
35	72	38.7	540	20 AAB59197	Human mammary sel-
36	72	38.7	540	22 AAB59200	Human mammary sel-
37	72	38.7	545	20 AAY22464	Human mammary sel-
38	72	38.7	545	22 AAB59196	Human mammary sel-
39	72	38.7	553	20 AAY22463	Human mammary sel-
40	72	38.7	553	22 AAB93475	Human mammary sel-
41	72	38.7	553	22 AAY22467	Human mammary sel-
42	72	38.7	559	20 AAB59199	Human mammary sel-
43	72	38.7	559	22 AAY22466	Human mammary sel-
44	72	38.7	589	20 AAB01204	Human GTPase assoc
45	72	38.7	589	21 AAB01204	Human GTPase assoc

## ALIGNMENTS

RESULT 1	AAAR85852	standard: peptide; 517 AA.
XX	AAAR85852:	
AC	13-SEP-1996 (first entry)	
XX	13-SEP-1996 (first entry)	
DT	WD-40 domain-contg. beta-TRCP protein.	
XX	WD-40 domain-contg. beta-TRCP protein.	
DE	WD40 repeat region: beta-transducin; protein-protein interaction; drug;	
XX	intracellular signalling; protein kinase C; homology; motif; modulator;	
KW	receptors of activated protein kinase; enzyme activity; isozyme; human.	
KW		
XX	Synthetic.	
OS	W09521252-A2.	
XX	10-AUG-1995.	
PN	31-JAN-1995: 95MO-US01210.	
XX	01-FEB-1994: 94US-0190802.	
XX	(STRD ) UNIV LELAND STANFORD JUNIOR.	
PA	Mochnly-Rosen D, Ron D:	
XX	Mochnly-Rosen D, Ron D:	
PI	WPI: 1995-283772/37.	
XX	New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the	
DR	activity of a protein, eg. protein kinase C, which interacts with a	
XX	protein contg. a WD-40 region.	
PT		

XX Example 5: Page 80-82; 351pp; English.  
 PS  
 CC Proteins AAR85851-92 are protein which contain at least one WD-40 (also  
 CC called beta-transducin homologous) amino acid repeat motifs. The WD-40  
 CC regions are involved in protein-protein interactions between proteins  
 CC involved in intracellular signalling. An example of such an interaction  
 CC is between protein kinase C and receptors of activated protein kinase  
 CC (RACK), esp. RACK-1 (AAR85850). Proteins AAR85851-82 were isolated based  
 CC on homology with beta-transducin, whereas proteins AAR85882-92 were  
 CC isolated based on homology with the WD-40 consensus sequence (AAR85893).  
 CC The proteins were used to construct the peptides AAR84928-R85063 and  
 CC AAR85786-R85842. The peptides can be used to identify target proteins  
 CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of  
 CC proteins involved in protein-protein interaction and to screen for drugs  
 CC that will affect protein-protein interaction involving WD-40 domains.

SQ Sequence 517 AA:

Query Match 100.0%; Score 186; DB 16; Length 517;  
 Best Local Similarity 100.0%; Pred. No. 3e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 34  
 |||||  
 DB 224 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 257

RESULT 2

ID AAY24054 standard; Protein: 569 AA.

XX AAY24054;

DT 30-SEP-1999 (first entry)

XX A human beta-transducin repeat containing protein.

XX Beta-transducin repeat containing protein; beta-Trcp; Skp1p;  
 KW Proteasome degradation pathway; Vpu protein; beta-catenin;  
 KW human immune deficiency virus-1; HIV-1; cellular protein; Ikappab;  
 KW ubiquitinylation; phosphorylated protein; tumour; apoptosis; Alzheimer's;  
 KW antiviral; antitumour; cell cycle regulation; protein degradation;  
 KW and anti-inflammatory; osteo-articular inflammation; acute inflammation;  
 KW tumour necrosis factor.

XX Homo sapiens.

XX Location/Qualifiers  
 FT 147..191

FT /note= "F box sequence"

FT 259..292

FT /note= "WD motif"

FT 304..332

FT /note= "WD motif"

FT 343..372

FT /note= "WD motif"

FT 387..415

FT /note= "WD motif"

FT 427..455

FT /note= "WD motif"

FT 467..492

FT /note= "WD motif"

FT 516..544

FT /note= "WD motif"

XX MO9938969-A1.

XX 05-AUG-1999.

XX 29-JAN-1999; 99WO-FR00196.

XX 09-DEC-1998; 98FR-0015545.

PR 30-JAN-1998; 98FR-0001100.

PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

PA (INSP ) INST PASTEUR.

PI Azenana Seisdedos F, Benarous R, Concordet J, Durand H;

PI Kroll M, Margotlin F,

DR WPI: 1999-469329/39.

DR N-PSDB; AAX86501.

PT New human beta-transducin repeat containing protein and its

PT fragments useful as, or to screen for, antiviral, antitumour,

PT anti-inflammatory and anti-Alzheimer's agents

PS Claim 1: Page 60-61; 71pp; French.

CC The present sequence represents a human beta-transducin repeat containing  
 CC protein (beta-Trcp). The protein directs proteins to the proteasome  
 CC degradation pathways. The protein is able to interact with the Vpu  
 CC protein of human immune deficiency virus-1 (HIV-1), cellular proteins  
 CC Ikappab or beta-catenin (bc) and/or protein Skp1p. The protein controls  
 CC ubiquitinylation of phosphorylated proteins and thus their targeting to  
 CC proteasomes for degradation. Depending on whether the process is  
 CC inhibited or promoted, the result may be delayed breakdown of CD4 (in  
 CC cases of HIV-1 infection); increased activity of Ikb (and thus reduced  
 CC activity of NFkappab) and increased degradation of mutant bc in tumour  
 CC cells, or increased bc survival (and reduced apoptosis) in Alzheimer's  
 CC patients. The beta-Trcp protein, and its active peptide fragments, or its  
 CC nucleic acid, are used to screen for anti HIV-1 agents (antivirals),  
 CC antitumour agents that disrupt cell cycle regulation or protein  
 CC degradation in human tumour cells, and anti-inflammatory agents that  
 CC disrupt activation by NFkappab. Fragments of the protein are also  
 CC useful for treating osteo-articular inflammation or acute inflammation  
 CC associated with release of tumour necrosis factor.

SQ Sequence 569 AA:

Query Match 100.0%; Score 186; DB 20; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 34  
 |||||  
 DB 259 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 292

RESULT 3

ID AAB12812 standard; protein: 569 AA.

XX AAB12812;

DT 27-NOV-2000 (first entry)

XX Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.

XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;  
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KW beta-Trcp.

XX Mus musculus.

OS JP2000166542-A.

PN 20-JUN-2000.

XX 02-DEC-1998; 98JP-0343437.

XX 02-DEC-1998; 98JP-0343437.

XX 02-DEC-1998; 98JP-0343437.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JTGODAN.

XX

PA

XX WPI: 2000-485550/43.  
 DR N-PSDB: AAA73131.  
 XX F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of IkappaB or beta-catenin  
 XX  
 PS Claim 2; Page 9-10; 19pp; Japanese.  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and Wp40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 CC  
 SQ Sequence 569 AA:  
 Query Match 100.0%; Score 186; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RHCRSETSKGYCLOYDOKIVSGLRDNTIKIW 34  
 ||||||||||||||||||||||||||||  
 Db 259 RHCRSETSKGYCLOYDOKIVSGLRDNTIKIW 292  
 RESULT 4  
 AAB12813 standard; protein; 569 AA.  
 XX AAB12813:  
 XX 27-NOV-2000 (first entry)  
 DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.  
 XX  
 KM ubiqlitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;  
 KM beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;  
 KM gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KM beta-TrCP.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP2000166542-A.  
 PD 20-JUN-2000.  
 XX  
 PF 02-DEC-1998; 98JP-0343437.  
 XX  
 PR 02-DEC-1998; 98JP-0343437.  
 XX  
 PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 XX  
 DR WPI: 2000-485550/43.  
 DR N-PSDB: AAA73132.  
 XX  
 PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of IkappaB or beta-catenin  
 XX  
 PS Claim 3; Page 10-12; 19pp; Japanese.  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and Wp40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for

CC the gene therapy of colon cancer by being recombined to a virus vector.  
 CC  
 SQ Sequence 569 AA:  
 Query Match 100.0%; Score 186; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RHCRSETSKGYCLOYDOKIVSGLRDNTIKIW 34  
 ||||||||||||||||||||||||||||  
 Db 259 RHCRSETSKGYCLOYDOKIVSGLRDNTIKIW 292  
 RESULT 5  
 AAY96697 standard; protein; 569 AA.  
 XX AAY96697:  
 XX 26-SEP-2000 (first entry)  
 DE Human beta-TrCP.  
 XX  
 KM E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KM nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KM anti-inflammatory; immunosuppressive; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200034447-A2.  
 PD 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99WO-US29371.  
 XX  
 PR 10-DEC-1998; 98US-0210060.  
 XX  
 PA (SIGN-) SIGNAL PHARM INC.  
 PA (YISS) YISSOM RES & DEV CO.  
 XX  
 PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;  
 PI Lavon I, Yaron A;  
 DR WPI: 2000-431294/37.  
 DR N-PSDB: AAA51229.  
 XX  
 PT Polypeptide enhancing phosphorylated IkappaB ubiquitination useful for  
 PT treating disorder associated with NF-kappaB activation e.g. cancer,  
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
 PT variant  
 XX  
 PS Claim 21; Page 72-74; 77pp; English.  
 CC Human beta-TrCP, an F-box/WD protein family member, has been shown to  
 CC have homology to human E3 ubiquitin ligase (E3). E3 enhances  
 CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of  
 CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B  
 CC degradation via the ubiquitin pathway is useful for identifying  
 CC modulators of this process for use in treating diseases associated with  
 CC activation of NF-kappa-B. In vitro analysis suggests that deletion of  
 CC the F-box results in a protein that functions as a dominant negative  
 CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a  
 CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha  
 CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated  
 CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B  
 CC inflammatory diseases, autoimmune diseases, cancer and viral infections.  
 CC  
 SQ Sequence 569 AA:  
 Query Match 100.0%; Score 186; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 34  
 Db 259 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 292

RESULT 6  
 AA83041

ID AAY83041 standard; Protein: 569 AA.

AC AAY83041;

DT 16-AUG-2000 (first entry)

DE F-box protein FBP-1.

XX F-box protein FBP-1.

KW F-box protein; FBP: diagnosis; treatment; screening; agonist;

KW antagonist; proliferative disorder; differentiative disorder;

KW breast cancer; prostate cancer; ovarian cancer; cancer;

KW small cell lung carcinoma; immune disorder; cardiovascular disorder;

XX inflammatory disorder; human.

OS Homo sapiens.

PN MO200012679-A1.

PD 09-MAR-2000.

PF 27-AUG-1999; 99WO-US19560.

PR 28-AUG-1998; 98US-0098355.

PR 03-FEB-1999; 98US-0118568.

PR 15-MAR-1999; 99US-0124449.

XX (UNYV ) UNIV NEW YORK STATE.

PA Chlaur DS, Pagano M, Latres E;

PI WPI: 2000-256635/22.

DR N-PSDB: AA293350.

XX Novel nucleic acid for screening compounds useful for treating

PT proliferative and differentiative disorders such as cancer and immune

PT disorders comprises sequences encoding ubiquitin ligases

PS Disclosure: Figure 3a; 245pp; English.

XX Nucleic acids encoding substrate-targeting subunits of ubiquitin

CC ligases with F-box motifs (F-box proteins) are useful for diagnosis

CC of proliferative and differentiated related disorders by measuring

CC FBP gene expression. Cells expressing such proteins or

CC their fragments are useful for screening compounds. The compounds

CC are agonists or antagonists, which are useful for treating a

CC breast, ovarian and prostate cancer and small cell lung carcinoma

CC and also major opportunistic infections, immune disorders,

CC cardiovascular diseases and inflammatory disorders. FBP protein,

CC analogs, derivatives and their subsequences, anti-FBP antibodies

CC are also useful in diagnosis of the disorders.

XX Sequence 569 AA;

QY Query Match 100.0%; Score 186; DB 21; Length 569;

Db Best Local Similarity 100.0%; Pred. No. 3.4e-18;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 34

Db 259 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 292

ID AAY83250 standard; Protein: 569 AA.

AC AAY83250;

DT 16-AUG-2000 (first entry)

DE F-box protein hbeatrCp.

XX Ubiquitin ligase; SCF: F-box protein; targeted degradation;

KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;

KW oncoprotein; Huntington's disease; gene knockout; delivery systems;

XX human.

OS Homo sapiens.

PN WO200022110-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US23705.

PR 09-OCT-1998; 98US-0103787.

PA (HARD ) HARVARD COLLEGE.

PI Zhou P, Howley P;

DR WPI: 2000-317970/27.

DR N-PSDB: AA293710.

XX Targeting degradation of polypeptide useful for treating cancer and

PT other proliferative disorders, involves conjugating polypeptide with

PT ubiquitin protein ligase or inhibiting ubiquitination using organic

PT compound

XX Claim 9; Page 171; 185pp; English.

CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin

CC ligases) which can be used for the targeted degradation of a target

CC polypeptide in vivo. Targeted degradation is achieved by expressing

CC the ubiquitin ligase in a cell linked to the interaction domain of

CC the target polypeptide and thereby recruiting the target polypeptide

CC to the ubiquitin ligase. Such methods are useful for decreasing or

CC increasing the level of a target polypeptide and for creating and

CC expressing a destabilized polypeptide which is subjected to SCF

CC mediated proteolysis. Degrading any desired protein in a cell is

CC useful for preventing or treating diseases caused by the presence of

CC abnormal amount of the specific polypeptides, for drug discovery and

CC for gene therapy. Diseases treated include cancer, by degradation of

CC oncoproteins, Huntington's disease, other proliferative disorders and

CC microbial infections. The method provides a quick and easy

CC alternative to gene knockout technology. The target polypeptide can

CC be degraded at all stages, or a specific stage, of development in the

CC mature animal.

XX Sequence 569 AA;

QY Query Match 100.0%; Score 186; DB 21; Length 569;

Db Best Local Similarity 100.0%; Pred. No. 3.4e-18;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 34

Db 259 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 292

ID AAY83254 standard; Protein: 569 AA.

AC AAY83254;

DT 16-AUG-2000 (first entry)

DE F-box protein hbeatrCp.

XX Ubiquitin ligase; SCF: F-box protein; targeted degradation;

KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;

KW oncoprotein; Huntington's disease; gene knockout; delivery systems;

XX human.

OS Homo sapiens.

PN WO200022110-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US23705.

PR 09-OCT-1998; 98US-0103787.

PA (HARD ) HARVARD COLLEGE.

PI Zhou P, Howley P;

DR WPI: 2000-317970/27.

DR N-PSDB: AA293710.

XX Targeting degradation of polypeptide useful for treating cancer and

PT other proliferative disorders, involves conjugating polypeptide with

PT ubiquitin protein ligase or inhibiting ubiquitination using organic

PT compound

XX Claim 9; Page 171; 185pp; English.

CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin

CC ligases) which can be used for the targeted degradation of a target

CC polypeptide in vivo. Targeted degradation is achieved by expressing

CC the ubiquitin ligase in a cell linked to the interaction domain of

CC the target polypeptide and thereby recruiting the target polypeptide

CC to the ubiquitin ligase. Such methods are useful for decreasing or

CC increasing the level of a target polypeptide and for creating and

CC expressing a destabilized polypeptide which is subjected to SCF

CC mediated proteolysis. Degrading any desired protein in a cell is

CC useful for preventing or treating diseases caused by the presence of

CC abnormal amount of the specific polypeptides, for drug discovery and

CC for gene therapy. Diseases treated include cancer, by degradation of

CC oncoproteins, Huntington's disease, other proliferative disorders and

CC microbial infections. The method provides a quick and easy

CC alternative to gene knockout technology. The target polypeptide can

CC be degraded at all stages, or a specific stage, of development in the

CC mature animal.

XX Sequence 569 AA;

QY Query Match 100.0%; Score 186; DB 21; Length 569;

Db Best Local Similarity 100.0%; Pred. No. 3.4e-18;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 34

Db 259 RHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 292



DR N-PSDB: AAZ29233.  
 XX  
 CC Human cell signalling proteins useful for, e.g. diagnosing cell  
 PT proliferative and inflammatory disorders  
 PS Claim 1; Page 77-78; 90pp; English.  
 XX  
 CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded  
 CC by cDNA obtained from Incyte clone 3239149 of COLAUC201 library. It is  
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is  
 CC found to be homologous to beta-transducin repeats containing  
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as  
 CC hybridisation probe for detecting CSIGP related sequences or allelic  
 CC variants. Recombinant CSIGP can be produced in host cells by transforming  
 CC them with genetically engineered vectors. Agonists or antagonists can be  
 CC used in the treatment of cell proliferative and inflammatory disorders  
 CC associated with decreased or increased CSIGP expression. CSIGP is used in  
 CC the diagnosis, prevention and treatment of cell proliferative disorders  
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.  
 XX  
 SQ Sequence 569 AA:

Query Match 100.0%; Score 186; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTTKIW 34  
 ||||||||||||||||||||||||||||||||||||  
 DB 259 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTTKIW 292

RESULT 10  
 AAB48298  
 ID AAB48298 standard; Protein: 569 AA.  
 XX  
 AC AAB48298;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE Human ZF11 protein.

XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; 2F;  
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
 KW Bad; Bcl-2; tumour; cytosolic.  
 XX  
 OS Homo sapiens.  
 PN WO200075184-A1.  
 XX  
 PD 14-DEC-2000.  
 XX  
 PF 05-JUN-2000; 2000WO-US15449.  
 XX  
 PR 04-JUN-1999; 9905-0137494.  
 XX  
 PA (UYVA ) UNTV YALE.  
 XX  
 PI Zhang H, Tsvetkov IM, Kondo T;  
 XX  
 DR WPI: 2001-061703/07.  
 DR N-PSDB; AAC84610.  
 XX  
 XX

Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
 PT involves altering levels of proteins such as S-phase kinase associated  
 XX proteins 1, 2 and cullin/CDC53 proteins -  
 XX

Claim 3; Page 130-132; 162pp; English.

The invention relates to methods of altering the polypeptide levels in a  
 CC cell, using proteins selected from S-phase kinase associated proteins 1  
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
 CC cullin/CDC53 family of proteins). The method is useful for altering the

CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for  
 CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
 CC that modulate interactions between SKP and target proteins are useful for  
 CC treating tumours.  
 XX  
 SQ Sequence 569 AA:

Query Match 100.0%; Score 186; DB 22; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3,4e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTTKIW 34  
 ||||||||||||||||||||||||||||||||||||  
 DB 259 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTTKIW 292

RESULT 11  
 AAM78583  
 ID AAM78583 standard; Protein: 579 AA.  
 XX  
 AC AAM78583;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human protein SEQ ID NO 1245.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 PN WO200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI: 2001-476283/51.  
 DR N-PSDB; AAK51716.  
 XX  
 XX

Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX

Claim 20; Page 3504-3505; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and

CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM00847) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX Sequence 579 AA;

Query Match 100.0%; Score 186; DB 22; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 34  
 |||||  
 Db 269 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 302

RESULT 12

AAM00847 ID AAM00847 standard; Protein; 590 AA.

XX AAM00847;

DT 01-OCT-2001 (first entry)

XX Human bone marrow protein, SEQ ID NO: 210.

XX Human; bone marrow; antineoplastic; cytostatic; neuroprotective;

KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;

KW immunosuppressive; gene therapy; cytokine cell proliferation;

KW cell differentiation modulator; immune disorder; infection; cancer;

KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.

OS Homo sapiens.

XX WO200153453-A2.

XX 26-JUL-2001.

XX 23-DEC-2000; 2000WO-US34960.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0634550.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 30-NOV-2000; 2000US-0250583.

XX (HYSE-) HYSEQ INC.

XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;

XX Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Drmanac RT;

XX WPI: 2001-488707/53.

XX N-P-SDB; AAKH9966.

XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
 for treating e.g. cancer and immune deficiency disorders -  
 Claim 10; Page 354-355; 648pp; English.

CC encodes is useful for the diagnosis and/or prognosis of one  
 CC or more types of cancer. The polynucleotide and polypeptide can be  
 CC used as nutritional sources or supplements and in the screening of  
 CC chemical compounds as potential drugs.

XX Sequence 590 AA;

Query Match 100.0%; Score 186; DB 22; Length 590;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 34  
 |||||  
 Db 280 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIM 313

RESULT 13

AAM78582 ID AAM78582 standard; Protein; 605 AA.

XX AAM78582;

DT 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1244.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX Zhao QA, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J, Ren F, Chen R, Wang ZW;

XX WPI: 2001-476283/51.

XX N-P-SDB; AAK51715.

XX Nucleic acids encoding polypeptides with cytokine-like activities,  
 useful in diagnosis and gene therapy -  
 Claim 20; Page 3503-3504; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX

SO Sequence 605 AA;

Query Match 100.0%; Score 186; DB 22; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 3,7e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCILOYDDOKIVSGLRDNTIKIW 34  
 Db 295 RHCRSETSKGVYCILOYDDOKIVSGLRDNTIKIW 328

## RESULT 14

AAAM00960  
 ID AAM00960 standard; Protein; 608 AA.

XX AAM00960;

DT 01-OCT-2001 (first entry)

DE Human bone marrow protein, SEQ ID NO: 436.

XX Human: bone marrow; antiinflammatory; cytostatic; neuroprotective;  
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;  
 KW immunosuppressive; gene therapy; cytokine cell proliferation;  
 KW cell differentiation modulator; immune disorder; infection; cancer;  
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.  
 XX  
 OS Homo sapiens.

PN WO200153453-A2.

PD 26-JUL-2001.

PF 23-DEC-2000; 2000WO-US34960.

PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 30-NOV-2000; 2000US-0250583.

XX (HYSE-) HYSEQ INC.

PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
 PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang Y;  
 PI Zhao QA, Zhou P, Drmanac RT;  
 XX  
 DR WPI: 2001-488707/53.  
 DR N-PSDB: AAK90079.

PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
 PT for treating e.g. cancer and immune deficiency disorders -

XX Claim 10; Page 523-524; 648pp; English.

XX The present sequence is one of 251 novel human polypeptides encoded  
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the  
 CC polypeptide encoded by it are useful in the treatment of various  
 CC immune deficiencies and disorders. The deficiencies and disorders may  
 CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal  
 CC infection, or may result from an autoimmune disorder, a coagulation  
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
 CC suppression of an inflammatory response or treatment of a nervous  
 CC system disorder such as Alzheimer's disease. Detection of the presence  
 CC or increased expression of the polynucleotide or the protein it  
 CC encodes is useful for the diagnosis and/or prognosis of one

CC or more types of cancer. The polynucleotide and polypeptide can be  
 CC used as nutritional sources or supplements and in the screening of  
 CC chemical compounds as potential drugs.  
 XX

SO Sequence 608 AA;

Query Match 100.0%; Score 186; DB 22; Length 608;  
 Best Local Similarity 100.0%; Pred. No. 3,7e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCILOYDDOKIVSGLRDNTIKIW 34  
 Db 298 RHCRSETSKGVYCILOYDDOKIVSGLRDNTIKIW 331

## RESULT 15

AAAM78584  
 ID AAM78584 standard; Protein; 632 AA.

XX AAM78584;

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1246.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX

OS Homo sapiens.

PN WO200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US04098.

PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI: 2001-476283/51.  
 DR N-PSDB: AAK51717.

PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -

XX Claim 20; Page 3505-3507; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666



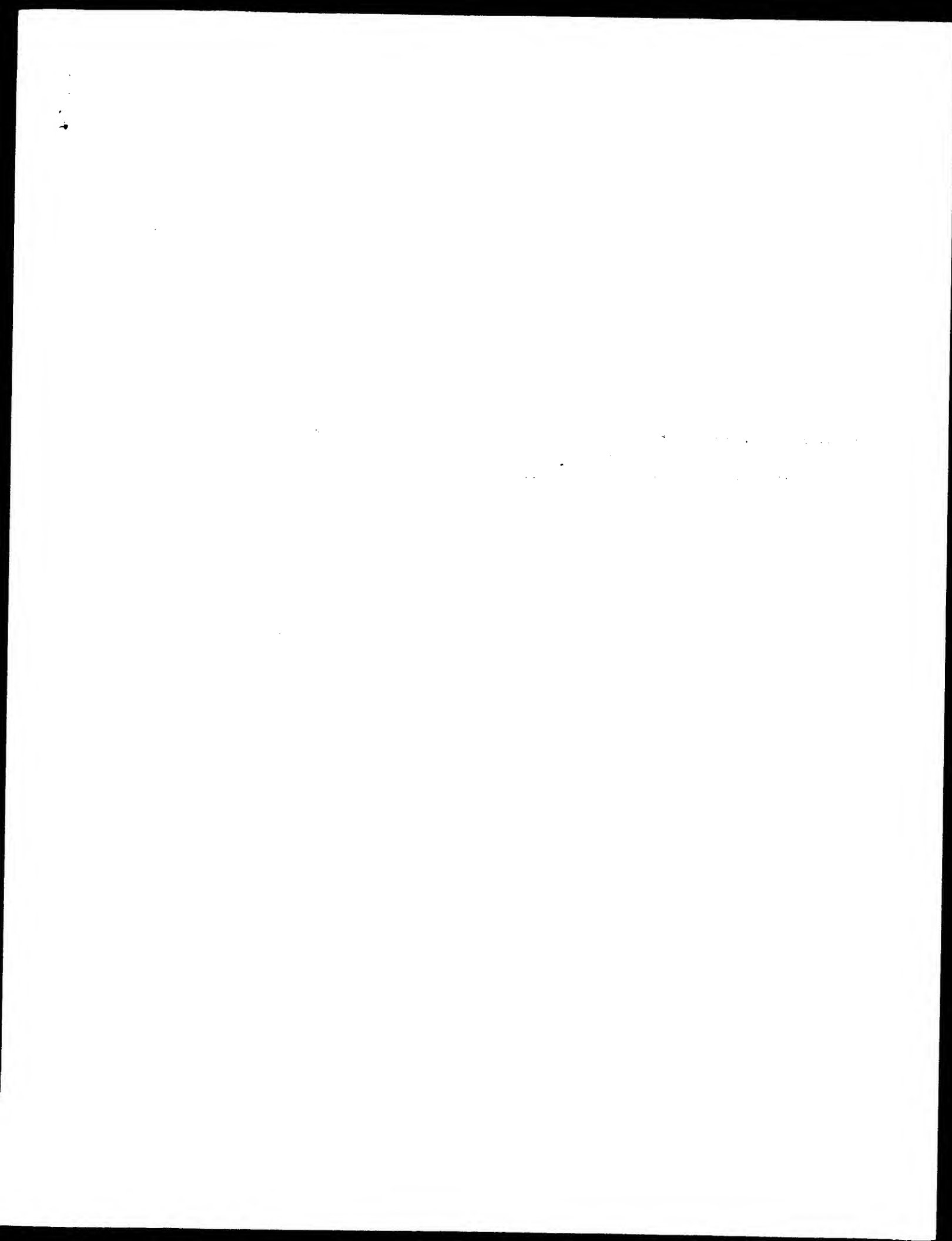
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX  
 SQ Sequence 632 AA:

Query Match 100.0%; Score 186; DB 22; Length 632;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGYCLOYDDOKIVSGLRDNTKIM 34  
 ||||||||||||||||||||||||||||||||  
 DB 322 RHCRSETSKGYCLOYDDOKIVSGLRDNTKIM 355

Search completed: April 11, 2003, 11:48:13  
 Job time : 11.4775 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 3.5201 Seconds  
(without alignments)  
284.191 Million cell updates/sec

Title: US-09-601-168B-2\_COPY\_259\_292

Perfect score: 186  
Sequence: 1 RHCRESKGVYCLQYDDQKIVSGLRDNTIKIW 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/3A COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	517	1 US-08-190-802A-30	Sequence 30, Appl
2	186	100.0	517	4 US-08-477-346-30	Sequence 30, Appl
3	186	100.0	517	4 US-08-473-089-30	Sequence 30, Appl
4	186	100.0	517	4 US-08-487-072A-30	Sequence 30, Appl
5	181	97.3	34	1 US-08-190-802A-83	Sequence 83, Appl
6	181	97.3	34	4 US-08-477-346-83	Sequence 83, Appl
7	181	97.3	34	4 US-08-473-089-83	Sequence 83, Appl
8	181	97.3	34	4 US-08-487-072A-83	Sequence 83, Appl
9	87	45.8	29	1 US-08-190-802A-87	Sequence 87, Appl
10	87	45.8	29	4 US-08-477-346-87	Sequence 87, Appl
11	87	45.8	29	4 US-08-473-089-87	Sequence 87, Appl
12	87	45.8	29	4 US-08-487-072A-87	Sequence 87, Appl
13	73	39.2	640	4 US-09-177-165A-30	Sequence 88, Appl
14	70	37.6	29	1 US-08-190-802A-88	Sequence 88, Appl
15	70	37.6	29	4 US-08-477-346-88	Sequence 88, Appl
16	70	37.6	29	4 US-08-473-089-88	Sequence 88, Appl
17	70	37.6	29	4 US-08-487-072A-88	Sequence 88, Appl
18	68.5	36.8	28	1 US-08-190-802A-84	Sequence 84, Appl
19	68.5	36.8	28	4 US-08-477-346-84	Sequence 84, Appl
20	68.5	36.8	28	4 US-08-473-089-84	Sequence 84, Appl
21	68.5	36.8	28	4 US-08-487-072A-84	Sequence 84, Appl
22	66	35.5	209	3 US-08-899-578-6	Sequence 6, Appl
23	66	35.5	587	3 US-08-899-578-2	Sequence 2, Appl
24	64	34.4	29	1 US-08-190-802A-89	Sequence 89, Appl
25	64	34.4	29	4 US-08-477-346-89	Sequence 89, Appl
26	64	34.4	29	4 US-08-473-089-89	Sequence 89, Appl
27	64	34.4	29	4 US-08-487-072A-89	Sequence 89, Appl

28	63	33.9	31	1 US-08-190-802A-71	Sequence 71, Appl
29	63	33.9	31	1 US-08-190-802A-102	Sequence 102, App
30	63	33.9	31	1 US-08-190-802A-130	Sequence 130, App
31	63	33.9	31	1 US-08-190-802A-174	Sequence 174, App
32	63	33.9	31	4 US-08-477-346-71	Sequence 71, Appl
33	63	33.9	31	4 US-08-477-346-102	Sequence 102, App
34	63	33.9	31	4 US-08-477-346-130	Sequence 130, App
35	63	33.9	31	4 US-08-477-346-174	Sequence 174, App
36	63	33.9	31	4 US-08-473-089-71	Sequence 71, Appl
37	63	33.9	31	4 US-08-473-089-102	Sequence 102, App
38	63	33.9	31	4 US-08-473-089-130	Sequence 130, App
39	63	33.9	31	4 US-08-473-089-174	Sequence 174, App
40	63	33.9	31	4 US-08-487-072A-71	Sequence 71, Appl
41	63	33.9	31	4 US-08-487-072A-102	Sequence 102, App
42	63	33.9	31	4 US-08-487-072A-130	Sequence 130, App
43	63	33.9	31	4 US-08-487-072A-174	Sequence 174, App
44	63	33.9	317	1 US-08-190-802A-27	Sequence 27, Appl
45	63	33.9	317	1 US-08-190-802A-41	Sequence 41, Appl

## ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
Sequence 30, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theroof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Delhinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R. 33,875  
REGISTRATION/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-190-802A-30

Query Match 100.0%; Score 186; DB 1; Length 517;  
Best Local Similarity 100.0%; Pred. No. 5.6e-19;  
Matches 34; Conservative 0; Mismatches 0; Gaps 0;

OY 1 RHCRESKGVYCLQYDDQKIVSGLRDNTIKIW 34

Db 224 RHCRSETSKGVCLQYDDQKIVSGLRDNTIKIM 257

## RESULT 2

US-08-477-346-30

Sequence 30, Application US/08477346  
Patent No. 6262023

## GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theeof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison &amp; Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477.346

FILING DATE: 07-JUN-1995

PRIORITY APPLICATION DATA:

CLASSIFICATION: 514

APPLICATION NUMBER: 08/487,072

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: MORASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: Peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13

US-08-477-346-30

Query Match

Best Local Similarity 100.0%; Score 186; DB 4; Length 517;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 RHCRSETSKGVCLQYDDQKIVSGLRDNTIKIM 257

## RESULT 3

US-08-473-089-30

Sequence 30, Application US/08473089

Patent No. 6342368

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theeof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison &amp; Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473.089

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MORASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.22

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids

TYPE: amino acid

TOPOLOGY: unknown

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13

US-08-473-089-30

Query Match

Best Local Similarity 100.0%; Score 186; DB 4; Length 517;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 RHCRSETSKGVCLQYDDQKIVSGLRDNTIKIM 257

Db 224 RHCRSETSKGVCLQYDDQKIVSGLRDNTIKIM 257

## RESULT 4

US-08-487-072A-30

Sequence 30, Application US/08487072A

Patent No. 6423684

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theeof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison &amp; Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MORASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-487-072A-30

Query Match 100.0%; Score 186; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 5.6e-19;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 34  
DB 224 IHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 257

RESULT 5  
US-08-190-802A-83  
Sequence 83, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Delinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 510  
ATTORNEY/AGENT INFORMATION:  
NAME: Fadian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-190-802A-83

Query Match 97.3%; Score 181; DB 1; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 34

DB 1 IHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 33

RESULT 6  
US-08-477-346-83  
Sequence 83, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHICE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-477-346-83

Query Match 97.3%; Score 181; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 34  
DB 1 IHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIW 33

RESULT 7  
US-08-473-089-83  
Sequence 83, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473, 089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-473-089-83  
Query Match  
Best Local Similarity 97.3%; Score 181; DB 4; Length 34;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 IHCRTSKGVYCYDQDKIVSGLRDNTIKIM 34  
DB 1 IHCRTSKGVYCYDQDKIVSGLRDNTIKIM 33  
RESULT 8  
US-08-487-072A-83  
Sequence 83, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487, 072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-487-072A-83  
Query Match  
Best Local Similarity 97.3%; Score 181; DB 4; Length 34;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 IHCRTSKGVYCYDQDKIVSGLRDNTIKIM 34  
DB 1 IHCRTSKGVYCYDQDKIVSGLRDNTIKIM 33  
RESULT 9  
US-08-190-802A-87  
Sequence 87, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190, 802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rV, Fig. 13  
US-08-190-802A-87  
Query Match  
Best Local Similarity 46.8%; Score 87; DB 1; Length 29;  
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
QY 10 KGVYCYDQDKIVSGLRDNTIKIM 34

Fri Apr 11 13:31:08 2003

us-09-601-168b-2\_copy\_259\_292.rai

Page 5

Db 4 RGIACLOYRDRLVSGSSDNTIRLM 28

RESULT 10

US-08-477-346-87

Sequence 87, Application US/08477346

Patent No. 6262023

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morlison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,346

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,072

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: MORASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13

US-08-477-346-87

Query Match 46.8%; Score 87; DB 4; Length 29;

Best Local Similarity 56.0%; Pred. No. 4, 1e-06;

Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 10 KGVYCLQYDDQKIYSGIRDNTIKIM 34

DB 4 RGIACLOYRDRLVSGSSDNTIRLM 28

RESULT 11

US-08-473-089-87

Sequence 87, Application US/08473089

Patent No. 6342368

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morlison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MORASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13

US-08-473-089-87

Query Match 46.8%; Score 87; DB 4; Length 29;

Best Local Similarity 56.0%; Pred. No. 4, 1e-06;

Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 10 KGVYCLQYDDQKIYSGIRDNTIKIM 34

DB 4 RGIACLOYRDRLVSGSSDNTIRLM 28

RESULT 12

US-08-487-072A-87

Sequence 87, Application US/08487072A

Patent No. 6423684

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morlison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MORASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13

US-08-473-089-87

Query Match 46.8%; Score 87; DB 4; Length 29;

Best Local Similarity 56.0%; Pred. No. 4, 1e-06;

Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 10 KGVYCLQYDDQKIYSGIRDNTIKIM 34

DB 4 RGIACLOYRDRLVSGSSDNTIRLM 28

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,089

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MORASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.22

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13

US-08-473-089-87

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVI, Fig. 13  
US-08-487-072A-87

Query Match 46.8%; Score 87; DB 4; Length 29;  
Best Local Similarity 56.0%; Pred. No. 4,1e-06;  
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 10 KGVYCIQYDDOKIVSGLRNTIKIW 34  
DB 4 RGACIQYRDLRVSSGSDNTIRLW 28

RESULT 13  
US-09-177-165A-30  
Sequence 30, Application US/09177165A  
Patent No. 6426205  
GENERAL INFORMATION:  
APPLICANT: Tyers, Mike  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN  
FILE REFERENCE: 11757.10USU1  
CURRENT FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: 60/092,443  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/063,254  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 30  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-177-165A-30

Query Match 39.2%; Score 73; DB 4; Length 640;  
Best Local Similarity 46.2%; Pred. No. 0.018;  
Matches 12; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 9 SKGYCCIQYDDOKIVSGLRNTIKIW 34  
DB 342 SDGVKLYFDKRLITGSLDKTIRW 367

RESULT 14  
US-08-190-802A-88  
Sequence 88, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theoreof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESS: Dellinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fadian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVI, Fig. 13  
US-08-190-802A-88

Query Match 37.6%; Score 70; DB 1; Length 29;  
Best Local Similarity 47.8%; Pred. No. 0.0012;  
Matches 11; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 12 VYCIQYDDOKIVSGLRNTIKIW 34  
DB 6 VCIQYDDOKIVSGLRNTIKIW 28

RESULT 15  
US-08-477-346-88  
Sequence 88, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theoreof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESS: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763

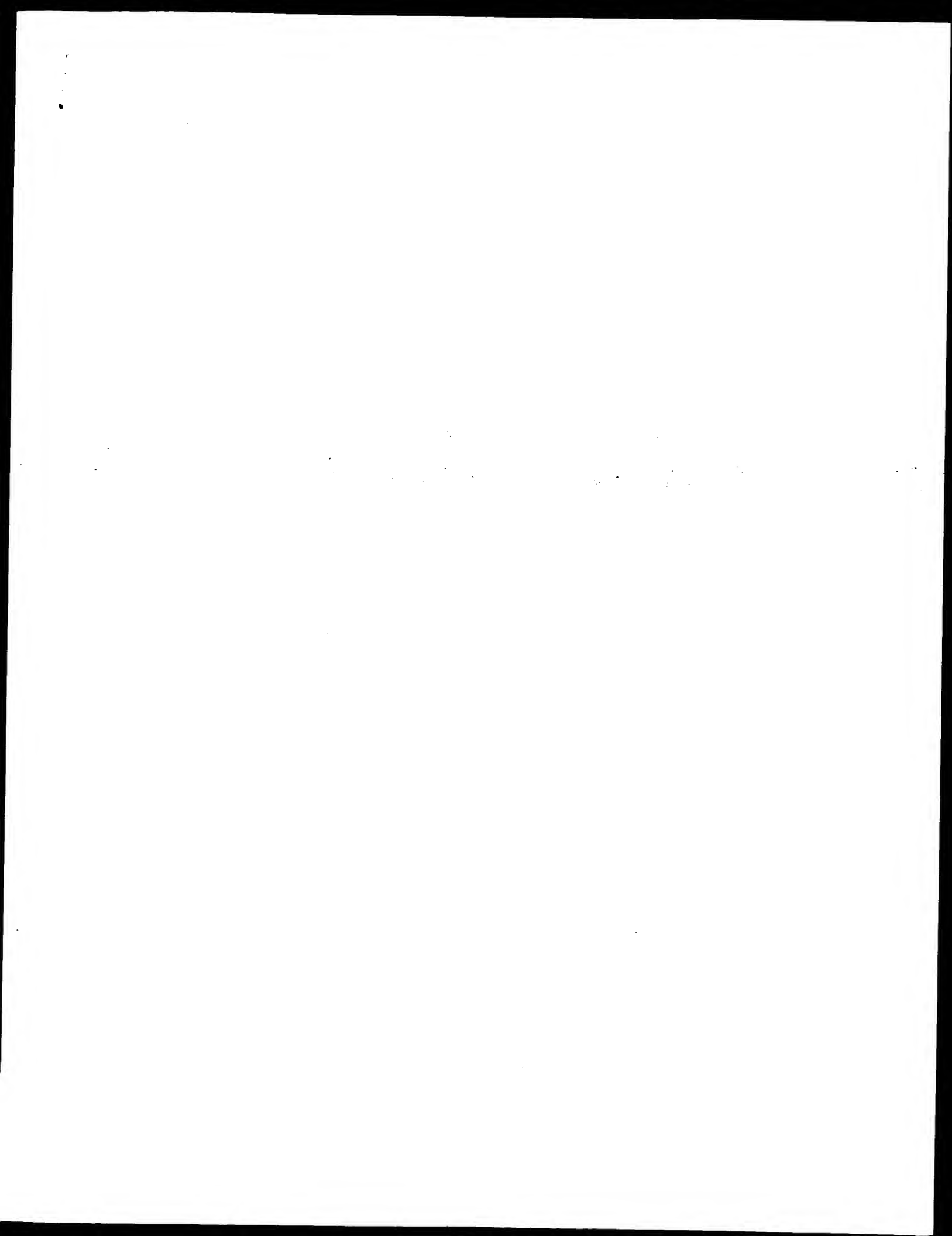


```

: INFORMATION FOR SEQ ID NO: 88:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 29 amino acids
:   TYPE: amino acid
:   TOPOLOGY: unknown
:   MOLECULE TYPE: peptide
:   HYPOTHETICAL: NO
:   ANTI-SENSE: NO
:   ORIGINAL SOURCE:
:   INDIVIDUAL ISOLATE: BETA TRCP IVI, Fig. 13
: US-08-477-346-88
:
: Query Match      37.6%; Score 70; DB 4; Length 29;
: Best Local Similarity 47.8%; Pred. No. 0.0012;
: Matches 11; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
:
: QY      12 VYCLQYDDQKIVSGLRDNTIKIW 34
:         | |::|::| | | | | | | |
: Db       6 VRCIRFDNKRIVSGAYDGKIKIW 28

```

Search completed: April 10, 2003, 12:51:17  
 Job time : 3.5201 secs



Fri Apr 11 13:31:08 2003

us-09-601-168b-2\_copy\_259\_292.rapb

Page 1

GenCore version 5.1.4.p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:49:06 ; Search time 3.47868 Seconds  
597.532 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_259\_292  
Perfect score: 186  
Sequence: 1 RHCRSEKSGVYCYQYDDQKIVSGLRDNTIKIW 34

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications, AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	186	100.0	569	US-10-038-010-8	Sequence 8, Appli
2	186	100.0	569	US-10-042-417-2	Sequence 2, Appli
3	123	66.1	265	US-10-116-016-30	Sequence 30, Appli
4	123	66.1	265	US-09-764-848-30	Sequence 30, Appli
5	90	48.4	678	US-09-801-368-314	Sequence 314, App
6	73	39.2	640	US-10-060-019-30	Sequence 30, Appli
7	72	38.7	540	US-09-213-888-7	Sequence 7, Appli
8	72	38.7	540	US-09-213-888-10	Sequence 10, Appli
9	72	38.7	540	US-09-328-877A-7	Sequence 7, Appli
10	72	38.7	540	US-09-328-877A-10	Sequence 10, Appli
11	72	38.7	545	US-09-213-888-6	Sequence 6, Appli
12	72	38.7	545	US-09-328-877A-6	Sequence 6, Appli
13	72	38.7	553	US-09-213-888-5	Sequence 5, Appli
14	72	38.7	553	US-09-328-877A-5	Sequence 5, Appli
15	72	38.7	559	US-09-213-888-9	Sequence 9, Appli
16	72	38.7	559	US-09-328-877A-9	Sequence 9, Appli
17	72	38.7	589	US-09-213-888-8	Sequence 8, Appli
18	72	38.7	589	US-09-328-877A-8	Sequence 8, Appli
19	72	38.7	592	US-09-213-888-4	Sequence 4, Appli

20	72	38.7	592	US-09-328-877A-4	Sequence 4, Appli
21	72	38.7	626	US-09-213-888-21	Sequence 21, Appli
22	72	38.7	626	US-09-328-877A-21	Sequence 21, Appli
23	72	38.7	627	US-09-213-888-3	Sequence 3, Appli
24	72	38.7	627	US-09-328-877A-3	Sequence 3, Appli
25	72	38.7	666	US-09-213-888-27	Sequence 27, Appli
26	72	38.7	666	US-09-328-877A-27	Sequence 27, Appli
27	72	38.7	669	US-09-213-888-25	Sequence 25, Appli
28	72	38.7	669	US-09-328-877A-25	Sequence 25, Appli
29	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
30	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
31	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
32	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
33	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
34	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
35	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
36	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
37	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
38	63	33.9	114	US-09-796-692-1541	Sequence 1541, Ap
39	59.5	32.0	190	US-10-016-447-4	Sequence 4, Appli
40	59	31.7	114	US-09-796-692-1541	Sequence 1541, Ap
41	59	31.7	296	US-10-083-357-1300	Sequence 1300, Ap
42	59	31.7	296	US-10-083-357-1300	Sequence 1300, Ap
43	58	31.2	1146	US-09-994-485-6	Sequence 6, Appli
44	58	31.2	1146	US-09-994-485-6	Sequence 6, Appli
45	57	30.6	64	US-09-843-843-20	Sequence 20, Appli

#### ALIGNMENTS

RESULT 1  
US-10-038-010-8  
Sequence 8, Application US/10038010  
Publication No. US2003040089A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS  
APPLICANT: Pierre, Legrain  
TITLE OF INVENTION: Protein-protein interactions in adipocyte cells  
FILE REFERENCE: B47674  
CURRENT APPLICATION NUMBER: US/10/038, 010  
CURRENT FILING DATE: 2002-07-23  
PRIOR APPLICATION NUMBER: US 60/259, 377  
PRIOR FILING DATE: 2001-01-02  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 8  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: beta-TCP1  
LOCATION: (1)-(569)  
OTHER INFORMATION:  
US-10-038-010-8

Query Match 100.0%; Score 186; DB 9; Length 569;  
Best Local Similarity 100.0%; Pred. No. 2.2e-18;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSEKSGVYCYQYDDQKIVSGLRDNTIKIW 34  
Db 259 RHCRSEKSGVYCYQYDDQKIVSGLRDNTIKIW 292

RESULT 2  
US-10-042-417-2  
Sequence 2, Application US/10042417  
Patent No. US20020123082A1  
GENERAL INFORMATION:  
APPLICANT: Paganon, M.  
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS

FILE REFERENCE: 5914-090-999  
CURRENT APPLICATION NUMBER: US/10/042,417  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-042-417-2

Query Match 100.0%; Score 186; DB 12; Length 569;  
Best Local Similarity 100.0%; Pred. No. 2,226; 18;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIW 34  
DB 259 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIW 292

RESULT 3  
US-10-116-016-30  
Sequence 30, Application US/10116016  
Publication No. US20030054379A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: Ptz08C1  
CURRENT APPLICATION NUMBER: US/10/116,016  
CURRENT FILING DATE: 2002-04-05  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 265  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-116-016-30

Query Match 66.1%; Score 123; DB 9; Length 265;  
Best Local Similarity 84.6%; Pred. No. 8e-10;  
Matches 22; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDQKIVSGLR 26  
DB 225 RHCRSETSKGVYCLQYDDQKIVSGLR 250

RESULT 4  
US-09-764-848-30  
Sequence 30, Application US/09764848  
Patent No. US20020077270A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: Ptz08  
CURRENT APPLICATION NUMBER: US/09/764,848  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 265  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-848-30

Query Match 66.1%; Score 123; DB 10; Length 265;  
Best Local Similarity 84.6%; Pred. No. 8e-10;  
Matches 22; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RHCRSETSKGVYCLQYDDQKIVSGLR 26  
DB 225 RHCRSETSKGVYCLQYDDQKIVSGLR 250

RESULT 5  
US-09-801-368-314  
Sequence 314, Application US/09801368  
Patent No. US20020128250A1  
GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Call, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 314  
LENGTH: 678  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
US-09-801-368-314

Query Match 48.4%; Score 90; DB 10; Length 678;  
Best Local Similarity 58.3%; Pred. No. 0.00011;  
Matches 14; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 11 GVCYCLQYDDQKIVSGLRDNTIKIW 34  
DB 391 GVCYCLQYDDQKIVSGLRDNTIKIW 414

RESULT 6  
US-10-060-019-30  
Sequence 30, Application US/10060019  
Publication No. US20030003564A1  
GENERAL INFORMATION:  
APPLICANT: Tyers, Mike  
APPLICANT: Williams, Andrew  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN  
FILE REFERENCE: 11757.100S01  
CURRENT APPLICATION NUMBER: US/10/060,019  
CURRENT FILING DATE: 2002-01-29  
PRIOR APPLICATION NUMBER: US/09/177,165  
PRIOR FILING DATE: 1998-10-22  
Prior Application removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 443  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/063,254  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 30  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-060-019-30

Query Match 39.2% Score 73; DB 9; Length 640;  
Best Local Similarity 46.2%; Pred. No. 0.028;  
Matches 12; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 9 SKGYCLOYDDOKIVSGLRDNTIKIW 34  
DB 342 SDGVKTYLFPDDRKLITGSLDKRTIRW 367

RESULT 7  
US-09-213-888-7  
Sequence 7, Application US/09213888A  
Patent No. US20020164683A1

GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-7

Query Match 38.7% Score 72; DB 9; Length 540;  
Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 12 VYCLQYDDOKIVSGLRDNTIKIW 34  
DB 217 ITCLOFCGNRIVSGSDNTLKW 239

RESULT 8  
US-09-213-888-10  
Sequence 10, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 10  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-10

Query Match 38.7% Score 72; DB 9; Length 540;  
Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 12 VYCLQYDDOKIVSGLRDNTIKIW 34  
DB 217 ITCLOFCGNRIVSGSDNTLKW 239

RESULT 9

US-09-328-877A-7  
Sequence 7, Application US/09328877A  
Patent No. US20020177187A1

GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-7

Query Match 38.7% Score 72; DB 9; Length 540;  
Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 12 VYCLQYDDOKIVSGLRDNTIKIW 34  
DB 217 ITCLOFCGNRIVSGSDNTLKW 239

RESULT 10  
US-09-328-877A-10  
Sequence 10, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 10  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-10

Query Match 38.7% Score 72; DB 9; Length 540;  
Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 12 VYCLQYDDOKIVSGLRDNTIKIW 34  
DB 217 ITCLOFCGNRIVSGSDNTLKW 239

RESULT 11  
US-09-213-888-6  
Sequence 6, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 545  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-6

Query Match 38.7%; Score 72; DB 9; Length 545;  
Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIKIW 34  
: |||: |||: |||: |||:  
DB 222 ITCLQFCGNRIIVSGSDNTLTKW 244

RESULT 12  
US-09-328-877A-6  
Sequence 6, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 545  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-6

Query Match 38.7%; Score 72; DB 9; Length 545;  
Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIKIW 34  
: |||: |||: |||: |||:  
DB 222 ITCLQFCGNRIIVSGSDNTLTKW 244

RESULT 13  
US-09-213-888-5  
Sequence 5, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-5

Query Match 38.7%; Score 72; DB 9; Length 553;

Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIKIW 34  
: |||: |||: |||: |||:  
DB 230 ITCLQFCGNRIIVSGSDNTLTKW 252

RESULT 14  
US-09-328-877A-5  
Sequence 5, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-5

Query Match 38.7%; Score 72; DB 9; Length 553;  
Best Local Similarity 52.2%; Pred. No. 0.032;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIKIW 34  
: |||: |||: |||: |||:  
DB 230 ITCLQFCGNRIIVSGSDNTLTKW 252

RESULT 15  
US-09-213-888-9  
Sequence 9, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 559  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-9

Query Match 38.7%; Score 72; DB 9; Length 559;  
Best Local Similarity 52.2%; Pred. No. 0.033;  
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIKIW 34  
: |||: |||: |||: |||:  
DB 236 ITCLQFCGNRIIVSGSDNTLTKW 258

Search completed: April 10, 2003, 13:16:43  
JDD time : 4.47868 secs

GenCore version 5.1.4.P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 3.93423 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_259\_292

Perfect score: 186  
Sequence: 1 RHCRSETSKGYCLOYDDQKIVSGLRNTIKIW 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR:73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	186	100.0	518	B48088	beta-transducin re
2	165	88.7	701	T16607	hypothetical prote
3	89	47.8	506	T50211	WD-repeat protein
4	85	45.7	605	T38932	probable sulfur me
5	82	44.1	650	T46660	sulfur controller-
6	73	39.2	640	S49932	MET30 protein - ye
7	70	37.6	659	S38108	hypothetical prote
8	66	35.5	579	T22703	hypothetical prote
9	64	34.4	310	T43158	probable GRP-bindi
10	64	34.4	314	T43299	probable GRP-bindi
11	63.5	34.1	589	AG2400	WD-repeat protein
12	63	33.9	316	S57839	CPC2 protein - Neu
13	63	33.9	317	A33928	GRP-binding protei
14	63	33.9	317	B33928	GRP-binding protei
15	63	33.9	317	S45054	GRP-binding regula
16	63	33.9	317	S38398	GRP-binding regula
17	63	33.9	317	A36986	activated protein
18	63	33.9	318	S11904	GRP-binding regula
19	63	33.9	349	AE2568	WD repeat protein
20	63	33.9	1356	T18521	beta transducin-II
21	62	33.3	267	S62507	hypothetical trp-a
22	62	33.3	422	A56640	CDCA repeat unit-c
23	62	33.3	1101	T26919	hypothetical prote
24	61	32.8	703	T43557	F-box/WD-repeat pr
25	61	32.8	906	S35342	Golgi-associated p
26	61	32.8	906	S35312	coatomer complex
27	60.5	32.5	438	T45823	hypothetical prote
28	60.5	32.5	473	T33805	hypothetical prote
29	60.5	32.5	558	T40651	pre-mrna splicing

30	59.5	32.0	586	2	T38992	MD-40 repeat regul
31	59	31.7	325	2	T23309	hypothetical prote
32	59	31.7	779	2	S56245	cell division cont
33	59	31.7	1000	2	T21970	hypothetical prote
34	58	31.2	532	1	B34104	protein-tyrosine k
35	58	31.2	920	2	C96831	hypothetical prote
36	58	31.2	926	2	G96563	probable coatomer
37	58	31.2	1146	2	A55532	myosin-heavy-chain
38	58	31.2	1893	2	T27762	hypothetical prote
39	57.5	30.9	336	2	T02300	hypothetical prote
40	57	30.6	290	2	T06784	GRP-binding regula
41	57	30.6	325	2	T09613	probable GRP-bindi
42	57	30.6	325	2	T09613	probable GRP-bindi
43	57	30.6	327	2	S48839	guanine nucleotide
44	57	30.6	334	2	T03764	protein RMD - rice
45	57	30.6	523	1	TVEVMT	protein-tyrosine k

## ALIGNMENTS

## RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-trp

C:Species: Xenopus laevis (African clawed frog)

C>Date: 26-May-1994 #sequence, revision 26-May-1994 #text-change 21-Jul-2000

C:Accession: B48088

R:Spewak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: A48088, MUID:9330289, PMID:8393141

A:Accession: B48088

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SP>

A:Cross-references: GB:M8268; NID:929542; PIDN:AAA02810.1; PID:9295543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <WDI>

Query Match 100.0%; Score 186; DB 2; Length 518;  
Best Local Similarity 100.0%; Pred. No. 5.9e-18;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGYCLOYDDQKIVSGLRNTIKIW 34  
DB 224 RHCRSETSKGYCLOYDDQKIVSGLRNTIKIW 257

## RESULT 2

T16607 hypothetical protein K10B2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence, revision 20-Sep-1999 #text-change 20-Sep-1999

C:Accession: T16607

R:Miller, N.

Submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid K10B2.

A:Reference number: Z18545

A:Accession: T16607

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-701 <ML>

A:Cross-references: EMBL:U28730; NID:9860694; PID:9860695; PIDN:AAA6258.1; CESP:K10B

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K10B2.1

A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 88.7%; Score 165; DB 2; Length 701;  
Best Local Similarity 88.2%; Pred. No. 6.8e-15;  
Matches 30; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 RHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIW 34
    ||::|| ||||| ||||| ||||| |||||
Db 250 RINCQSENSKGVYCLQYDDDKIVSGLRDNTIKIW 283

```

### RESULT 3

120211  
 WP:repeat protein [imported] - fission yeast (*Schizosaccharomyces pombe*)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000  
 C:Accession: T50211  
 R:McDougal1, R.C.; Rajadream, M.A.; Barreil1, B.G.; Brown, S.; Murphy, L.; Jones, L.; Mc  
 submitted to the EMBL Data Library, January 2000  
 A:Reference number: 225046  
 A:Accession: T50211  
 A:Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-506 <MCDD>  
 A:Cross-references: EMBL:AL136538; PUDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05  
 A:Experimental source: strain 972h(-); cosmid C30  
 C:Genetics:  
 A:Gene: SPAC29F6.01; SPDB:SPAC30.05  
 A:Map position: 1  
 A:Introns: 43/1; 74/3

```

Query Match 47.8% Score 89; DB 2; Length 506;
Best Local Similarity 51.9%; Pred. No. 0.00019;
Matches 14; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY      8 TSKGYCGLDYPDQKIVSGLRDNTIKIW 34
      1 | | | | | : | | | | : | | : |
Db 220 TLDSTVCVQVDDKIMVSSGKDPFVSW 246

```

RESULT 5  
T46660  
sulfur controller-2 protein [imported] - *Neurospora crassa*  
C:Species: *Neurospora crassa*  
C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
R:Accession: T46660  
R:Kumar, A.; Paietta, J.V.  
Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995

A:Title: The sulfur controller-2 negative regulatory gene of *Neurospora crassa* encode  
A:Reference number: 223121; MUID:95541499; PMID:7724564  
A:Accession: T46660  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-650 <K0M>  
A:Cross-References: EMBL:U17251; NID:9806757; PIDN:AA68968.1; PID:9806758  
C:Genetics:  
A:Gene: scon-2  
A:Map position: 3  
A:Introns: 75/3; 319/1; 354/1  
C:Function:  
A:Description: negatively regulates sulfur structural gene expression  
A>Note: scon-2+ expression is dependent on CYS3 function and the binding of CYS3 to t  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match	44.1%;	Score 82;	DB 2;	Length 650;
Best Local Similarity	54.2%;	Pred. No. 0.0023;		
Matches 13;	Conservative 6;	Mismatches 5;	Indels 0;	Gaps 0;

QY 11 GVCYCYDDOKIVSGIDNTIKIV 34  
I: ||:| | :||| | :|||:|  
Db 336 GIRALQFDSDSKISGSLDHTIKW 359

RESULT 6  
S49932

Multo proteoform - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: protein Y19005.02; protein Y1L046w  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 28-May-1993 #sequence  
 C:Accession: S49932; S43750  
 R:Orell, C.; Bowman, S  
 submitted to the EMBL Data Library, December 1994  
 A:Reference number: S49931  
 A:Accession: S49932  
 A:Molecule type: DNA  
 A:Residues: 1-640 <CODE>  
 A:Cross-references: GB:Z47047; EMBL:Z46861; NID:9603997; P1D:9763300; MIPS:Y1L046w  
 R:Thomas, D.; Cherest, H.; Barbey, R.; Studdin-Kerfjan, Y.  
 submitted to the EMBL Data Library, December 1993  
 A:Reference number: S43750  
 A:Accession: S43750  
 A:Molecule type: DNA  
 A:Residues: 1-60, I', 62-640 <THO>  
 A:Cross-references: EMBL:L26505; NID:g432493; P1D:g432494  
 C:Genetics:  
 A:Gene: SGD:MEF30  
 A:Cross-references: SGD:S0001308; MIPS:Y1L046w  
 A:Map position: 9L  
 F:298-329/Domain: WD repeat proteins; WD repeat homology  
 F:329-369/Domain: WD repeat homology <WD1>  
 F:374-409/Domain: WD repeat homology <WD2>  
 F:417-450/Domain: WD repeat homology <WD3>  
 F:417-450/Domain: WD repeat homology <WD4>

RESULT 7  
S38108  
hypochemical protein YK036c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-May-1994 #sequence\_revise 03-May-1994 #text\_change 08-Oct-1999  
C:Accession: S38108  
R:Urrutiarazu, L.A.; Janniaux, J.C.  
submitted to the Protein Sequence Database, March 1994





OY 6 SETSKGYCQYDDQKIVSGLRDNTIKIM 34  
 DB 345 TDVTRAI-ALTPDDQTLISGSADKTIKIM 372

## RESULT 12

S57839  
 C:Species: Neurospora crassa  
 C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 21-Jul-2000  
 C:Accession: S57839  
 R:Mueller, F.; Krueger, D.; Sattlegger, E.; Hoffmann, B.; Ballario, P.; Kanan, M.; Bart  
 Mol. Gen. Genet. 248, 162-173, 1995  
 A:Title: The cpc-2 gene of *Neurospora crassa* encodes a protein entirely composed of WD-  
 A:Reference number: S57839; MUID:95379761; PMID:7651339  
 A:Accession: S57839  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-316 <MUR>  
 A:Cross-references: EMBL:X81875; NID:971565; PIDN:CAA57460.1; PID:971566  
 C:Genetics:  
 A:Introns: 14/3; 27/2; 69/3; 175/3  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 F:11-45/Domain: WD repeat homology <WD1>  
 F:59-92/Domain: WD repeat homology <WD2>  
 F:101-134/Domain: WD repeat homology <WD3>  
 F:144-179/Domain: WD repeat homology <WD4>  
 F:188-221/Domain: WD repeat homology <WD5>  
 F:231-261/Domain: WD repeat homology <WD6>  
 F:279-312/Domain: WD repeat homology <WD7>

Query Match 33.9%; Score 63; DB 2; Length 316;  
 Best Local Similarity 64.7%; Pred. No. 0.48;  
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 18 DDQKIVSGLRDNTIKIM 34  
 DB 116 DNRQIVSGSRDRTIKIM 132

## RESULT 13

A33928  
 C:Species: Gallus gallus (chicken)  
 C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 21-Jul-2000  
 C:Accession: A33928  
 R:Guillemot, F.; Billault, A.; Auffray, C.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 4594-4598, 1989  
 A:Title: Physical linkage of a guanine nucleotide-binding protein-related gene to the ch  
 A:Reference number: A33928; MUID:89282817; PMID:2499885  
 A:Accession: A33928  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-317 <GUI>  
 A:Cross-references: GB:M24193; NID:9212293; PIDN:AAA50559.1; PID:9212294  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 F:11-45/Domain: WD repeat homology <WD1>  
 F:59-92/Domain: WD repeat homology <WD2>  
 F:101-134/Domain: WD repeat homology <WD3>  
 F:144-179/Domain: WD repeat homology <WD4>  
 F:188-221/Domain: WD repeat homology <WD5>  
 F:231-261/Domain: WD repeat homology <WD6>  
 F:279-312/Domain: WD repeat homology <WD7>

Query Match 33.9%; Score 63; DB 2; Length 317;  
 Best Local Similarity 64.7%; Pred. No. 0.48;  
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 18 DDQKIVSGLRDNTIKIM 34  
 DB 116 DNRQIVSGSRDRTIKIM 132

## RESULT 14

B33928  
 C:Species: Homo sapiens (man)  
 C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 26-May-2000  
 C:Accession: B33928  
 R:Guillemot, F.; Billault, A.; Auffray, C.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 4594-4598, 1989  
 A:Title: Physical linkage of a guanine nucleotide-binding protein-related gene to the  
 A:Reference number: A33928; MUID:89282817; PMID:2499885  
 A:Accession: B33928  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-317 <GUI>  
 A:Cross-references: GB:M24194; NID:9187701; PIDN:AAA59626.1; PID:9307218  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 F:11-45/Domain: WD repeat homology <WD1>  
 F:59-92/Domain: WD repeat homology <WD2>  
 F:101-134/Domain: WD repeat homology <WD3>  
 F:144-179/Domain: WD repeat homology <WD4>  
 F:188-221/Domain: WD repeat homology <WD5>  
 F:231-261/Domain: WD repeat homology <WD6>  
 F:279-312/Domain: WD repeat homology <WD7>

Query Match 33.9%; Score 63; DB 2; Length 317;  
 Best Local Similarity 64.7%; Pred. No. 0.48;  
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 18 DDQKIVSGLRDNTIKIM 34  
 DB 116 DNRQIVSGSRDRTIKIM 132

## RESULT 15

S45054  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 26-May-2000  
 C:Accession: S45054  
 R:Chou, Y.; Huang, J.; Liu, H.  
 submitted to the EMBL Data Library, May 1994  
 A:Description: cDNA of porcine G-beta like protein.  
 A:Reference number: S45054  
 A:Accession: S45054  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-317 <CHO>  
 A:Cross-references: EMBL:Z33879; NID:9495143; PIDN:CAA83944.1; PID:9495144  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 F:11-45/Domain: WD repeat homology <WD1>  
 F:59-92/Domain: WD repeat homology <WD2>  
 F:101-134/Domain: WD repeat homology <WD3>  
 F:144-179/Domain: WD repeat homology <WD4>  
 F:188-221/Domain: WD repeat homology <WD5>  
 F:231-261/Domain: WD repeat homology <WD6>  
 F:279-312/Domain: WD repeat homology <WD7>

Query Match 33.9%; Score 63; DB 2; Length 317;  
 Best Local Similarity 64.7%; Pred. No. 0.48;  
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 18 DDQKIVSGLRDNTIKIM 34  
 DB 116 DNRQIVSGSRDRTIKIM 132

Search completed: April 10, 2003, 13:23:00  
 Job time: 4.93423 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:51:26 ; Search time 2.44336 Seconds  
(without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_259\_292

Perfect score: 186  
Sequence: 1 RHCSESTSGVYCLQYDDQKIVSGLRDNTIKIW 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	186	100.0	518	1	TRCB_XENLA
2	186	100.0	605	1	FWIA_HUMAN
3	165	88.7	542	1	FWIB_HUMAN
4	165	88.7	665	1	LI23_CAEEL
5	90	48.4	678	1	SCOB_EMENT
6	89	47.8	506	1	POFB_SCHPO
7	85	45.7	605	1	POFI_SCHPO
8	82	44.1	650	1	SCO2_NEUCR
9	76	40.9	684	1	CC4_CANAL
10	73	39.2	640	1	MT30_YEAST
11	70	37.6	659	1	YK16_YEAST
12	69	37.1	317	1	GBLP_HYCAT
13	66	35.5	379	1	SE10_CAEEL
14	64	34.4	314	1	GBLP_SCHPO
15	63	33.9	316	1	GBLP_BIOGL
16	63	33.9	316	1	GBLP_NEUCR
17	63	33.9	317	1	GBLP_BRARE
18	63	33.9	317	1	GBLP_HUMAN
19	63	33.9	317	1	GBLP_ORENI
20	63	33.9	318	1	GBLP_CHRE
21	63	33.9	318	1	GBLP_DROME
22	63	33.9	1356	1	HET1_PODAN
23	62	33.3	422	1	FBW2_HUMAN
24	62	33.3	422	1	FBW2_MOUSE
25	61	32.8	332	1	GBLP_DICDI
26	61	32.8	703	1	POP2_SCHPO
27	61	32.8	904	1	COPP_RAT
28	61	32.8	905	1	COPP_BOVIN
29	61	32.8	905	1	COPP_HUMAN
30	60	32.3	905	1	COPP_MOUSE
31	60	32.3	732	1	KMHA_DICDI
32	60	32.3	914	1	COPP_DROME
33	59.5	32.0	586	1	TU12_SCHPO

34	59	31.7	324	1	GBLP_CAEEL	Q21215	caenorhabdi
35	59	31.7	779	1	CC4_YEAST	P07834	saccharomyc
36	59	31.7	1000	1	COPP_CAEEL	Q20168	caenorhabdi
37	58	31.2	318	1	GBLP_TRYBB	Q94775	trypanosoma
38	58	31.2	531	1	SRG2_XENLA	P13116	xenopus lae
39	58	31.2	682	1	TU01_KL0LA	P56094	kiuyveromyc
40	58	31.2	1146	1	KMHA_DICDI	P42527	dictyosteli
41	57	30.6	325	1	GBLP_MEDSA	Q24076	medicago sa
42	57	30.6	325	1	GBLP_SOYBN	Q39836	glycine max
43	57	30.6	327	1	GBLP_ARATH	Q24456	arabidopsis
44	57	30.6	327	1	GBLP_BRANA	Q39336	brassica na
45	57	30.6	334	1	GBLP_ORYSA	P49027	oryza sativ

## ALIGNMENTS

RESULT 1  
TRCB\_XENLA STANDARD: PRT; 518 AA.  
ID AC 091854: P70037; P70038:  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Beta-TTCP (Beta-transducin repeat-containing protein).  
GN FBXW1 OR FBRCF.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
OC Xenopodinae; Xenopus.  
ON NCBI\_TaxID=8355;  
RX MEDLINE=93330289; PubMed=8393141;  
RA Spevak W., Keiper B.D., Stralowa C., Castranon M.J.;  
RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in  
anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein  
with beta-transducin repeats."  
RL Mol. Cell. Biol. 13:4953-4966(1993).  
RN [2]  
RP SEQUENCE OF 302-518 FROM N.A.  
RX MEDLINE=97109804; PubMed=8952061;  
RA Hudson J.W., Alarcon V.B., Elinson R.P.;  
RT "Identification of new localized RNAs in the Xenopus oocyte by  
RT differential display PCR."  
RL Dev. Genet. 19:190-198(1996).  
CC - FUNCTION: Probably recognizes and binds to some phosphorylated  
proteins and promotes their ubiquitination and degradation.  
CC - SUBUNIT: PART OF A SCF (SKP1-CUL1N-F-BOX) PROTEIN LIGASE COMPLEX  
(BY SIMILARITY).  
CC - DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-  
MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC  
GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO  
NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR  
TADPOLE EMBRYO.  
CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
CC - SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
CC - SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL; M98268; AA02810.1; -  
DR EMBL; U63921; AAB49671.1; -  
DR EMBL; U63922; AAB49672.1; -  
DR InterPro: IPR001810; F-box.  
DR InterPro: IPR001680; WD40.  
DR Pfam; PF00400; WD40; 7.

DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR Prodom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ubl conjugation pathway; Repeat; WD repeat.  
 FT DOMAIN 119 157 F-BOX.  
 FT REPEAT 230 258 WD 1.  
 FT REPEAT 270 298 WD 2.  
 FT REPEAT 310 338 WD 3.  
 FT REPEAT 353 381 WD 4.  
 FT REPEAT 393 421 WD 5.  
 FT REPEAT 433 461 WD 6.  
 FT REPEAT 482 510 WD 7.  
 FT CONFLICT 302 304 GEM -> EFR (IN REF. 2).  
 FT CONFLICT 516 518 GLA -> AAH (IN REF. 2).  
 SQ SEQUENCE 518 AA: 59507 MW: 2A52EC19028127F3 CRC64;

Query Match 100.0%; Score 186; DB 1; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTIKIM 34  
 Db 224 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTIKIM 257

RESULT 2  
 ID FW1A\_HUMAN STANDARD; PRT; 605 AA.  
 AC 09Y297; 09Y213;  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TrCP)  
 DE (E3SK1kappaB) (pikappaBalpha-E3 receptor subunit).  
 GN FBXW1A OR FBW1A OR BTCP OR BTFC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=99075339; PubMed=9859996;  
 RA Varon A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,  
 RT "Identification of the receptor component of the ikappaBalpha-  
 RT ubiquitin ligase";  
 RL Nature 396:590-594(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Lymphoid;  
 RX MEDLINE=98325370; PubMed=9660940;  
 RA Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,  
 RA Thomas D., Strebel K., Benarous R.;  
 RT "A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu  
 RT connects CD4 to the ER degradation pathway through an F-box motif";  
 RL Mol. Cell 1:565-574(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Genetelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99145464; PubMed=9990852;  
 RA Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,

RA Harper J.W.;  
 RT "The SCF(beta-TrCP)-ubiquitin ligase complex associates specifically  
 RT with phosphorylated destruction motifs in I-kappa-B-alpha and  
 RT beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro";  
 RL Genes Dev. 13:270-283(1999).  
 CC - FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA  
 CC (PIKAPPAALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR  
 CC UBIQUITINATION AND DEGRADATION.  
 CC - SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC - SUBCELLULAR LOCATION: Cytoplasmic.  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC - SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF101784; AAD08702.1; -  
 DR EMBL: Y14153; CA74572.1; -  
 DR EMBL: AF129530; AAF04464.1; -  
 DR GENE: HGNC:1144; BTCP.  
 DR MIM: 603482; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF004400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR Prodom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 FT DOMAIN 190 228 F-BOX.  
 FT REPEAT 301 338 WD 1.  
 FT REPEAT 341 378 WD 2.  
 FT REPEAT 381 418 WD 3.  
 FT REPEAT 424 461 WD 4.  
 FT REPEAT 464 503 WD 5.  
 FT REPEAT 505 541 WD 6.  
 FT REPEAT 553 590 WD 7.  
 FT VARSPLIC 17 52 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 605 AA: 68866 MW: 4C67F3B7E400FD37 CRC64;

Query Match 100.0%; Score 186; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTIKIM 34  
 Db 295 RHCRSETSKGYCYCLOYDDQKIVSGLRDNTIKIM 328

RESULT 3  
 ID FW1B\_HUMAN STANDARD; PRT; 542 AA.  
 AC 09URB1; 09Y4C6; 09P2S8; 09P2S9;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TrCP2).  
 DE FBXW1B OR FBW1B OR BTCP2 OR KIAA0696.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20030505; PubMed=10531035;  
 RA Cenciarrelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RT Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Fetal lung;  
 RX MEDLINE=20160458; PubMed=10694485;  
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,  
 RA Katoh M.;  
 RT "Molecular cloning and genomic structure of the betatropin gene on  
 RT chromosome 5q35.1.";  
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A, B AND C (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF176022; AAF04528.1; -;  
 DR EMBL: AB033279; BAA92329.1; -;  
 DR EMBL: AB033280; BAA92330.1; -;  
 DR EMBL: AB033281; BAA92331.1; -;  
 DR EMBL: AB034596; BAA31671.1; ALT\_INIT.  
 DR Genew: HGNC:13607; FBXW1B.  
 DR MIM: 605651; -;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00330; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 FT DOMAIN 129 167 F-BOX.  
 FT REPEAT 238 275 WD 1.  
 FT REPEAT 278 315 WD 2.  
 FT REPEAT 318 355 WD 3.  
 FT REPEAT 361 398 WD 4.  
 FT REPEAT 401 440 WD 5.  
 FT REPEAT 442 478 WD 6.  
 FT REPEAT 490 527 WD 7.  
 FT VARSPLIC 16 49 MISSING (IN ISOFORM A).

FT VARSPLIC 16 48 CSVPRLMTLGCANLIVSMCALSCLOMPSVACL -> NTSV  
 FT MEDONEDSPKKTILM (IN ISOFORM B).  
 SQ SEQUENCE 542 AA: 62090 MW: 7CD40087EFAA55C8A CRC64;  
 Query Match 88.7%; Score 165; DB 1; Length 542;  
 Best Local Similarity 85.3%; Pred. No. 1,4e-15;  
 Matches 29; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 RICHRETSKGYVCLQYDDOKIVSGLRDITIKW 34  
 DB 232 RIQRENSKGYVCLQYDDOKIVSGLRDINSIKW 265  
 RESULT 4  
 LIT3 CAEEL STANDARD; PRT; 665 AA.  
 AC 009990; O9G6N6; 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein Lin-23.  
 GN LIN-23 OR K10B2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION: DEVELOPMENTAL STAGE, AND MUTAGENESIS.  
 RX MEDLINE=20515608; PubMed=11060233.  
 RA Kipreos E.T., Golbel S.P., Hedgecock E.M.;  
 RT "The Caenorhabditis elegans F-box/WD-repeat protein Lin-23 functions  
 RT to limit cell division during development.";  
 RL Development 127:5071-5082(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Miller N.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Functions cell autonomously to negatively regulate cell  
 CC cycle progression. Required to restrain cell proliferation in  
 CC response to developmental cues. Probably recognizes and binds to  
 CC some proteins and promotes their ubiquitination and degradation  
 CC (By similarity).  
 CC -1- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex  
 CC (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- DEVELOPMENTAL STAGE: Highest levels in embryos and adults, lowest  
 CC levels in larvae. Maternal expression results in high zygotic  
 CC levels.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO X.LAEVIS FBXW1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF275253; AAG28037.1; -;  
 DR EMBL: U28730; AA68258.2; -;  
 DR WormPep: K10B2.1; CE28600.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.



```

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB061694; BAB5543.1; -
DR EMBL: AL136538; CAB66464.1; -
DR EMBL: Z66525; CA91423.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00440; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 3.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 8.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Repeat: WD repeat.
KW DOMAIN
FT REPEAT 219 256 F-BOX.
FT REPEAT 259 298 WD 1.
FT REPEAT 301 338 WD 2.
FT REPEAT 345 386 WD 3.
FT REPEAT 388 426 WD 4.
FT REPEAT 427 464 WD 5.
FT REPEAT 468 505 WD 6.
FT REPEAT 506 545 WD 7.
SQ SEQUENCE 506 AA; 58257 MW; CEF34DAEFFPC2E10 CRC64;

Query Match 47.8%; Score 89; DB 1; Length 506;
Best Local Similarity 51.9%; Pred. No. 5.7e-05;
Matches 14; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 8 TSKGVYCLQYDDOKIVSGLRDNTIKIW 34
1 |||:||||:||||:||||:||||:
Db 220 TLDVYCVQYDDEIMVSGSKDRTVSWV 246

RESULT 7
POFI_SCHPO STANDARD; PRT; 605 AA.
ID POF1_SCHPO
AC P87053;
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE F-box/WD-repeat protein pof1 (Skp1-binding protein 1).
OS POF1 OR SBP1 OR SPAC57A10.05C.
GN Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Katayama S., Kitamura K., Toda T.;
RT "Systematic genome-wide analysis of F-box protein-encoding genes in
RT fission yeast."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA MEDLINE=21848401; PubMed=118593360;
RA Wood V., Gilliam R., Najandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

```

```

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
RA Welfens I., Vanstreels E., Rieger M., Scheffer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Mamhuth R., Punnett B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB032410; BAA84528.1; -
DR EMBL: Z94864; CAB08168.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00440; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 5.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Ubl conjugation, Repeat: WD repeat.
KW DOMAIN
FT REPEAT 107 153 F-BOX.
FT REPEAT 171 199 WD 1.
FT REPEAT 271 299 WD 2.
FT REPEAT 311 339 WD 3.
FT REPEAT 350 379 WD 4.
FT REPEAT 390 420 WD 5.
FT REPEAT 432 460 WD 6.
FT REPEAT 472 500 WD 7.
FT REPEAT 510 538 WD 8.
SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 45.7%; Score 85; DB 1; Length 605;
Best Local Similarity 57.7%; Pred. No. 0.00025;
Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 9 SKGVYCLQYDDOKIVSGLRDNTIKIW 34
1 |||:||||:||||:||||:||||:
Db 313 SSGVTCLQYDDCKRLKSGMDKTIKRW 338

```

```

RESULT 8
SC02_NEUCR STANDARD: PRT: 650 AA.
AC 001277;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sulfur controller-2 (SCONZ).
GN SCON-2.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RX MEDLINE=95241499; PubMed=7724564;
RA Kumar A., Paletta J.V.;
RT "The sulfur controller-2 negative regulatory gene of Neurospora
RT crassa encodes a protein with beta-transducin repeats."
RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995)
CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC -1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U17251; AAA68966.1;
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001880; WD40.
DR Pfam: PF004400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS00682; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Repeat; WD repeat.
FT DOMAIN 124 170
FT REPEAT 292 320 WD 1.
FT REPEAT 332 360 WD 2.
FT REPEAT 372 400 WD 3.
FT REPEAT 411 441 WD 4.
FT REPEAT 453 488 WD 5.
FT REPEAT 528 564 WD 6.
FT REPEAT 576 604 WD 7.
FT REPEAT 616 644 WD 8.
SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 44.1%; Score 82; DB 1; Length 650;
Best Local Similarity 54.2%; Pred. No. 0.00071;
Matches 13; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

```

AC P53699;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cell division control protein 4.
GN CDC4.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SGY126;
RA Shieh J.C., White A.M., Rosamond J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC VARIOUS ASPECTS OF SPOKULATION. REQUIRED FOR HTAI-HTBI LOCUS
CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X96763; CA65538.1;
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001880; WD40.
DR Pfam: PF004400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 1.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS00682; WD_REPEATS_2; 4.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT DOMAIN 212 258
FT REPEAT 322 351 WD 1.
FT REPEAT 363 391 WD 2.
FT REPEAT 403 431 WD 3.
FT REPEAT 442 468 WD 4 (POTENTIAL).
FT REPEAT 478 506 WD 5.
FT REPEAT 519 549 WD 6.
FT REPEAT 561 589 WD 7.
SQ SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;

Query Match 40.9%; Score 76; DB 1; Length 684;
Best Local Similarity 44.1%; Pred. No. 0.0052;
Matches 15; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

```

```

RESULT 9
CC4_CANAL STANDARD: PRT: 684 AA.
ID CC4_CANAL
DB 336 GIRALQFDDSKILSGSLDHTIKW 359
QY 11 GYVCIQYDDQKIYSGLRDNTIKW 34
1: ||:||||:|:|:|:|

```

```

RESULT 10
MT30_YEAST STANDARD: PRT: 640 AA.
AC P39014;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MET30 protein.

```





```

OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydrozoa; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Zurich;
RA Hassel M.E.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions. There are no restrictions on its
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X97800; CA66387.1; -
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR PRINTS; PRO0320; GPROTEINBRPT.
DR ProDom: PD000018; WD40; 5.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS00682; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT REPEAT 3 45 WD 1.
FT REPEAT 52 93 WD 2.
FT REPEAT 94 135 WD 3.
FT REPEAT 137 180 WD 4.
FT REPEAT 181 221 WD 5.
FT REPEAT 222 262 WD 6.
FT REPEAT 263 313 WD 7.
SQ SEQUENCE 317 AA; 35256 MW; 5D3087F3B50E8A4F CRC64;

Query Match 37.1%; Score 69; DB 1; Length 317;
Best Local Similarity 70.6%; Pred. No. 0.022;
Matches 12; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 18 DDQKIVSGIRNTIKIM 34
DB 116 DNRQIVSGSRNTIKIM 132

RESULT 13
SE10.CAEEL
AC 093794; STANDARD; PRT; 579 AA.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sel-10 protein.
GN SEL-10 OR P55B12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Rhabditinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sims M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC [2]
CC REVISIONS.
CC Jones S.J.M.;
CC Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z79757; CAB02129.1; -
DR WormPep; F55B12.3; CE16120.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PRO0320; GPROTEINBRPT.
DR ProDom: PD000018; WD40; 5.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 5.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT DOMAIN 113 159 F-BOX.
FT REPEAT 245 274 WD 1.
FT REPEAT 286 316 WD 2.
FT REPEAT 328 356 WD 3.
FT REPEAT 368 396 WD 4.
FT REPEAT 408 438 WD 5.
FT REPEAT 453 481 WD 6.
FT REPEAT 493 522 WD 7.
SQ SEQUENCE 579 AA; 64275 MW; 0647245AF5964663 CRC64;

Query Match 35.5%; Score 66; DB 1; Length 579;
Best Local Similarity 43.5%; Pred. No. 0.11;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 12 VYCLQDDQKIVSGIRNTIKIM 34
DB 251 ITCMQIHDDVLVTGSDNTIKIM 273

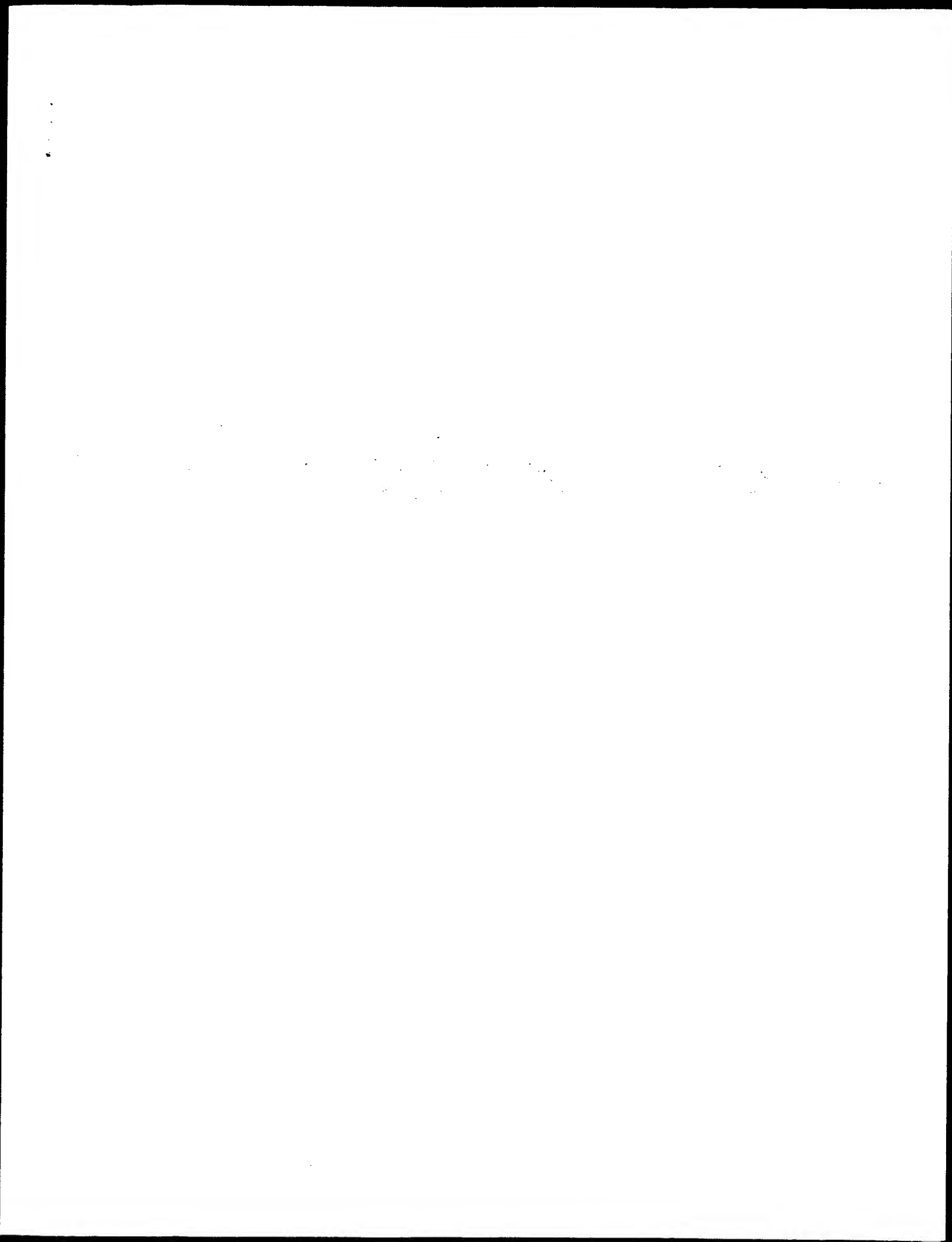
RESULT 14
GRLP_SCHPO
AC 010281; P78896; STANDARD; PRT; 314 AA.
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Guanine nucleotide-binding protein beta subunit-like protein (Receptor
DE of activated protein kinase C).
GN RKP1 OR CPC2 OR SPAC6B12.15.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ED616;
RA Park S.-K., Yoo H.-S.;
RA Park S.-K., Yoo H.-S.;
RA Kim H.-B., Yoo H.-S.;
RA "RKP1/Cpc2, a fission yeast RACK1 homolog, is involved in actin
RA cytoskeleton organization through protein kinase C, Pck2, signaling."
RL Biochem. Biophys. Res. Commun. 282:10-15(2001).
RN [3]

```

RP SEQUENCE FROM N.A.  
 RC SRRAIN-972:  
 RA MEDLINE-21848401; PubMed-11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Soutos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckart G., Aert R., Robben J., Gymnopoulos B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Beger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Sipakovski G.V., Ussey D., Barrell B.G., Nuse P.,  
 RA "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 RN [4]  
 RP SEQUENCE OF 5-314 FROM N.A.  
 RC STRAIN-PR745:  
 RX MEDLINE-98167722; PubMed-9501991;  
 RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.,  
 RT "Identification of open reading frames in Schizosaccharomyces pombe  
 RT cDNAs."  
 RL DNA Res. 4:363-369(1997).  
 CC -1- FUNCTION: MAY BE A RECEPTOR FOR PROTEIN KINASE C IN THE REGULATION  
 CC OF ACTIN CYTOSKELETON ORGANIZATION DURING CELL WALL SYNTHESIS AND  
 CC MORPHOGENESIS.  
 CC -1- SUBUNIT: INTERACTS WITH PKC2.  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH PARTICULATE FRACTIONS.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 DR EMBL: L37885; AAA56865.2; -;  
 DR EMBL: AF320333; AAK38633.1; -;  
 DR EMBL: Z98531; CAB11079.1; -;  
 DR EMBL: D89247; BAA13908.1; -;  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 5.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Repeat, WD repeat.  
 FT REPEAT 13 44 WD 1.  
 FT REPEAT 61 91 WD 2.  
 FT REPEAT 103 133 WD 3.  
 FT REPEAT 146 178 WD 4.  
 FT REPEAT 190 220 WD 5.  
 FT REPEAT 231 260 WD 6.  
 FT REPEAT 281 311 WD 7.  
 FT REPEAT 41 41 WD 7.  
 FT CONFLICT 41 41 I -> L (IN REF. 1).

SQ SEQUENCE 314 AA: 34851 MW: 4E14707164E68ACD CRC64;  
 Query Match 34.4%; Score 64; DB 1; Length 314;  
 Best Local Similarity 64.7%; Pred. No. 0.11;  
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 DDQKIVSGLRDNTIKIW 34  
 Db 116 DNRQVSGSRDNTIKIW 132  
 ::::|||||  
 RESULT 15  
 GBLP\_BIOGL  
 ID GBLP\_BIOGL STANDARD; PRT; 316 AA.  
 AC 093134;  
 DT 01-NOV-1997 (rel. 35. Created)  
 DT 01-NOV-1997 (rel. 35. Last sequence update)  
 DT 16-OCT-2001 (rel. 40. Last annotation update)  
 DE Guanine nucleotide-binding protein beta subunit-like protein  
 DE (Receptor of activated protein kinase C) (RACK).  
 OS Biophytaria glabrata (Bloodfluke planorb).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 OC Planorbidae; Biophytaria.  
 OX NCBI\_TaxID=6526;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lardans V.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 DR EMBL: U49437; AAB07039.1; -;  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 5.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Repeat, WD repeat.  
 FT REPEAT 4 46 WD 1.  
 FT REPEAT 52 93 WD 2.  
 FT REPEAT 94 135 WD 3.  
 FT REPEAT 137 180 WD 4.  
 FT REPEAT 181 221 WD 5.  
 FT REPEAT 222 263 WD 6.  
 FT REPEAT 264 312 WD 7.  
 SQ SEQUENCE 316 AA: 35008 MW: B668144BC7B81DB CRC64;  
 QY 18 DDQKIVSGLRDNTIKIW 34  
 Db 116 DNRQVSGSRDNTIKIW 132  
 ::::|||||  
 Query Match 33.9%; Score 63; DB 1; Length 316;  
 Best Local Similarity 64.7%; Pred. No. 0.15;  
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Search completed: April 10, 2003, 13:19:33  
 Job time : 3.44336 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 ; Search time 8.03411 Seconds

(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_259\_292

Sequence: 1 RHCRSEKSGVCLQYDDQKIVSGLRDNFIKIW 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21:\*

- 1: sp-archaea:\*
- 2: sp-bacteria:\*
- 3: sp-fungi:\*
- 4: sp-human:\*
- 5: sp-invertebrate:\*
- 6: sp-mammal:\*
- 7: sp-mhc:\*
- 8: sp-organelle:\*
- 9: sp-phage:\*
- 10: sp-plant:\*
- 11: sp-rodent:\*
- 12: sp-virus:\*
- 13: sp-vertebrate:\*
- 14: sp-unclassified:\*
- 15: sp-ryinus:\*
- 16: sp-bacteriap:\*
- 17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	569	11	Q9R1G7 mus musculus
2	186	100.0	569	11	Q9R1G7 mus musculus
3	186	100.0	569	11	Q9R1G7 mus musculus
4	171	91.9	47	13	Q8U0N5 mus musculus
5	167	89.8	510	5	Q44382 drosophila
6	167	89.8	510	5	Q9VDE3 drosophila
7	166	89.6	430	5	Q9H554 heterodera
8	163	87.6	252	11	Q9R23H0 mus musculus
9	87	46.8	252	11	Q9R23H0 mus musculus
10	75	40.3	1326	13	Q9V2F4 drosophila
11	72	38.7	122	13	Q8U0N3 mus musculus
12	72	38.7	553	4	Q9N0X6 homo sapien
13	72	38.7	561	4	Q96R12 homo sapien
14	72	38.7	589	4	Q961E0 homo sapien
15	72	38.7	627	4	Q96A16 homo sapien
16	72	38.7	629	11	Q8VHP4 mus musculus

17	72	38.7	629	11	Q8VHP4 mus musculus
18	72	38.7	707	4	Q969H0 homo sapien
19	67	36.0	942	5	Q96611 dictyostell
20	66	35.5	585	5	Q952T0 caenorhabdi
21	66	35.5	587	5	Q44083 caenorhabdi
22	64	34.4	781	5	Q8SVW7 encephalito
23	63.5	34.1	589	16	Q8YN14 anabaena sp
24	63	33.9	280	11	Q9CS00 mus musculus
25	63	33.9	296	11	Q9ERM6 mus musculus
26	63	33.9	313	5	Q9UAY2 euryyma sc
27	63	33.9	316	3	Q9HGV7 emericella
28	63	33.9	317	13	Q9W711 xenopus lae
29	63	33.9	319	5	Q9SPD5 heliothis v
30	63	33.9	349	16	Q8YJY6 anabaena sp
31	63	33.9	823	10	Q9FT96 arabidopsis
32	63	33.9	914	4	Q9C0P5 homo sapien
33	63	33.9	1065	3	Q8X1P6 podospora a
34	63	33.9	1356	3	Q8X1P5 podospora a
35	63	33.9	1356	3	Q8X1P4 podospora a
36	63	33.9	1356	3	Q8X1P3 podospora a
37	62	33.3	245	11	Q9CP06 mus musculus
38	62	33.3	245	11	Q99153 mus musculus
39	62	33.3	454	4	Q8W511 homo sapien
40	62	33.3	454	4	Q9HA09 homo sapien
41	62	33.3	454	11	Q9Q0H1 caenorhabdi
42	62	33.3	1101	5	Q62471 caenorhabdi
43	62	33.3	1573	11	Q90XL2 mus musculus
44	61	32.8	329	5	Q8SSP4 dictyostell
45	61	32.8	932	10	Q9FHL5 arabidopsis

## ALIGNMENTS

RESULT 1

Q9R1G7 PRELIMINARY: PRT: 569 AA.

AC Q9R1G7: 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Beta-transducin repeat-containing protein.

GN BTRC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RM [1]

RP SEQUENCE FROM N.A.

RA Winston J., Elledge S.J., Harper J.W.;

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

CC -I- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

DR EMBL: AF110396: AAP41025.1..

DR MGD: MGI:1338871: BTRC.

DR InterPro: IPR001810: WD-box.

DR InterPro: IPR001680: WD40.

DR Pfam: PF00646: F-box; 1.

DR Pfam: PF00400: WD40; 7.

DR PRINTS: PR00320: GPROTEINBPT.

DR ProDom: PD000018: WD40; 4.

DR SMART: SM00236: FBOX; 1.

DR SMART: SM00320: WD40; 7.

DR PROSITE: PS0181: FBOX; 1.

DR PROSITE: PS00678: WD\_REPEATS\_1; UNKNOWN\_6.

DR PROSITE: PS0082: WD\_REPEATS\_2; 7.

DR PROSITE: PS0294: WD\_REPEATS\_REGION; 1.

KW Repeat; WD repeat.

SEQUENCE 569 AA: 65209 MW; E6DDCAD28D551D9D CRC64;

Query Match 100.0%; Score 186; DB 11; Length 569;

Best Local Similarity 100.0%; Pred. No. 4,1e-18;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      1 RHCRSETSKGYCLOYDDOKIVSGLRDNTIKIM 34
      ||||||||||||||||||||||||||||||||
Db      259 RHCRSETSKGYCLOYDDOKIVSGLRDNTIKIM 292

RESULT 2
ID      092159      PRELIMINARY:      PRT:      569 AA.
AC      092159;
DT      01-MAY-1999 (TrEMBLrel. 10, Created)
DT      01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE      Beta-transducin repeat containing protein.
GN      BTRC.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99145465; PubMed=9990853;
RA      Spencer E., Jiang J., Chen Z.J.,
RT      "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT      Slim/Deaf-1".
CC      -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR      EMBL; AF112979; AAD04181.1; -.
DR      MGD; MGI:1338871; Btrc.
DR      InterPro; IPR001810; F-box.
DR      InterPro; IPR001680; WD40.
DR      Pfam; PF00646; F-box; 1.
DR      PRINTS; PR00320; GPROTEINRPT.
DR      ProDom; PD000018; WD40; 4.
DR      SMART; SM00256; FBOX; 1.
DR      SMART; SM00320; WD40; 7.
DR      PROSITE; PS50181; FBOX; 1.
DR      PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_6.
DR      PROSITE; PS50082; WD_REPEATS_2; 7.
DR      PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW      Repeat; WD repeat.
SQ      SEQUENCE 569 AA; 65047 MW; BC7C7444815BED96 CRC64;

Query Match      100.0%; Score 186; DB 11; Length 569;
Best Local Similarity 100.0%; Pred. No. 4,1e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 RHCRSETSKGYCLOYDDOKIVSGLRDNTIKIM 34
      ||||||||||||||||||||||||||||||||
Db      259 RHCRSETSKGYCLOYDDOKIVSGLRDNTIKIM 292

RESULT 3
ID      090015      PRELIMINARY:      PRT:      569 AA.
AC      090015;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE      Ubiquitin ligase FMD1 (Beta-transducin repeat containing protein)
DE      (F-box-WD40 repeat protein 1).
GN      BTRC OR FBXW1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99199275; PubMed=10097128;
RA      Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA      Hatori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA      Nakayama K.-I.;
RT      "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a

```

```

RT      ubiquitin ligase Skp1/Cul 1/F-box protein FMD1.";
RT      Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99075339; PubMed=9859996;
RA      Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA      Andersen J.S., Mann M., Mercutio F., Ben-Neriah Y.;
RT      "Identification of the receptor component of the IkappaBalpha-
RT      ubiquitin ligase.";
RT      Nature 396:590-594(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Strausberg R.;
RA      Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=129/SV;
RX      MEDLINE=21601157; PubMed=11735228;
RA      Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,
RA      Nakayama K.-I.;
RT      "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of
RT      Caenorhabditis elegans SEL-10.";
RT      Genomics 78:214-222(2001).
CC      -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR      EMBL; AF081887; AAD17755.1; -.
DR      EMBL; AF099932; AAD08701.1; -.
DR      EMBL; BC003989; AAL03989.1; -.
DR      EMBL; AF391190; AAL40929.1; -.
DR      EMBL; AF391179; AAL40929.1; JOINED.
DR      EMBL; AF391178; AAL40929.1; JOINED.
DR      EMBL; AF391180; AAL40929.1; JOINED.
DR      EMBL; AF391181; AAL40929.1; JOINED.
DR      EMBL; AF391182; AAL40929.1; JOINED.
DR      EMBL; AF391183; AAL40929.1; JOINED.
DR      EMBL; AF391184; AAL40929.1; JOINED.
DR      EMBL; AF391185; AAL40929.1; JOINED.
DR      EMBL; AF391186; AAL40929.1; JOINED.
DR      EMBL; AF391187; AAL40929.1; JOINED.
DR      EMBL; AF391188; AAL40929.1; JOINED.
DR      EMBL; AF391189; AAL40929.1; JOINED.
DR      MGD; MGI:1338871; Btrc.
DR      InterPro; IPR001810; F-box.
DR      InterPro; IPR001680; WD40.
DR      Pfam; PF00646; F-box; 1.
DR      Pfam; PF00400; WD40; 7.
DR      PRINTS; PR00320; GPROTEINRPT.
DR      ProDom; PD000018; WD40; 4.
DR      SMART; SM00256; FBOX; 1.
DR      SMART; SM00320; WD40; 7.
DR      PROSITE; PS50181; FBOX; 1.
DR      PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_6.
DR      PROSITE; PS50082; WD_REPEATS_2; 7.
DR      PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW      Ligase; Repeat; WD repeat.
SQ      SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Query Match      100.0%; Score 186; DB 11; Length 569;
Best Local Similarity 100.0%; Pred. No. 4,1e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 RHCRSETSKGYCLOYDDOKIVSGLRDNTIKIM 34
      ||||||||||||||||||||||||||||||||
Db      259 RHCRSETSKGYCLOYDDOKIVSGLRDNTIKIM 292

RESULT 4
ID      0800N5      PRELIMINARY:      PRT:      47 AA.
AC      0800N5;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE      Beta-Trcp protein (Fragment).

```

```

GN BETA-TRCP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Carnevali F.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ballarino M.;
RL "Analisi strutturale e funzionale del gene beta-TrCP in Xenopus laevis";
RT Thesis (2001);
RL Department of Genetics and Molecular Biology "Charles Darwin",
RL University of Rome La Sapienza, Rome, Italy.
DR EMBL: AJ428936; CAD21933.1; -.
FT NON_TER 1 47
FT SEQUENCE 47 AA; 5499 MW; 1C18C15B13EE9F5F CRC64;
SQ
Query Match 91.9%; Score 171; DB 13; Length 47;
Best local Similarity 100.0%; Pred. No. 4.5e-17;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHCSETSKGVYCLQYDDOKIVSGLRDNTIK 32
Db 15 RHCSETSKGVYCLQYDDOKIVSGLRDNTIK 46

RESULT 5
ID 044382 PRELIMINARY; PRT; 510 AA.
AC 044382;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE SLMB
GN SLMB OR SLMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98121115; PubMed=9461217;
RA Jiang J.; Struhl G.;
RL "Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slmb";
RT Nature 391:493-496(1998).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF032878; AAC38852.1; -.
DR FlyBase: FBgn0023423; slmb.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; DBB0243D3730A5E8 CRC64;

Query Match 89.8%; Score 167; DB 5; Length 510;
Best local Similarity 91.2%; Pred. No. 1.8e-15;
Matches 31; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 RHCSETSKGVYCLQYDDOKIVSGLRDNTIK 34
Db 202 RHCSETSKGVYCLQYDDOKIVSGLRDNTIK 235

RESULT 6
ID 09VDE3 PRELIMINARY; PRT; 510 AA.
AC 09VDE3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE SLMB protein (SLMB).
GN SLMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;
RA Amanatides P.G.; Scherer S.E.; Li P.W.; Hoskins R.A.; Galle R.F.;
RA George R.A.; Lewis S.E.; Richards S.; Ashburner M.; Henderson S.N.;
RA Sutton G.G.; Mortman J.R.; Yandell M.D.; Zhang Q.; Chen L.X.;
RA Brandon R.C.; Rogers Y.-H.C.; Blazek R.G.; Champe M.; Pfeiffer B.D.;
RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Miklos G.L.G.;
RA Abril J.F.; Agbayani A.; An H.-J.; Andrews-Pfannkoch C.; Baldwin D.;
RA Baller R.M.; Basu A.; Baxendale J.; Bayraktaroglu L.; Beasley E.M.;
RA Beeson K.Y.; Benos P.V.; Berman B.P.; Bhandari D.; Bolshakov S.;
RA Borkova D.; Botchan M.R.; Bouck J.; Brockstein P.; Brotlier P.;
RA Burtis K.C.; Busam D.A.; Butler H.; Cadieu E.; Center A.; Chandra I.;
RA Cherry J.M.; Cawley S.; Dahlke C.; Davenport L.B.; Davies P.;
RA de Pablo J.M.; Delcher A.; Deng Z.; Mays A.D.; Dew I.; Dietz S.M.;
RA Dodson K.; Doop L.E.; Downes M.; Dugan-Rocha S.; Dunkov B.C.; Dunn P.;
RA Durbin K.J.; Evangelista C.C.; Ferraz C.; Ferreira S.; Fleischmann W.;
RA Flosser C.; Gabriellian A.E.; Garg N.S.; Gelbart W.M.; Glasser K.;
RA Goddek A.; Gong F.; Gorrell J.H.; Gu Z.; Guan P.; Harris M.;
RA Harris N.L.; Harvey D.; Heiman T.J.; Hernandez J.R.; Houck J.;
RA Hostin D.; Houston K.A.; Howland T.J.; Wei M.-H.; Ibegwan C.;
RA Jalili M.; Kalush F.; Karpen G.H.; Ke Z.; Kennison J.A.; Ketchum K.A.;
RA Kimmel B.E.; Kodira C.D.; Kraft C.; Kravitz S.; Kulp D.; Lai Z.;
RA Lasko P.; Lei Y.; Levitsky A.A.; Li J.; Li Z.; Liang Y.; Lin X.;
RA Liu X.; Mattei B.; McIntosh T.C.; McLeod M.P.; McPherson D.;
RA Merkulov G.; Mishina N.V.; Mobarry C.; Morris J.; Moshrefi A.;
RA Mount S.M.; Moy M.; Murphy B.; Murphy L.; Muzny D.M.; Nelson D.L.;
RA Nelson D.R.; Nelson K.A.; Nixon K.; Nusskern D.R.; Paclet J.M.;
RA Palazzolo M.; Pittman G.S.; Pan S.; Pollard J.; Puri V.; Reese M.G.;
RA Reibert K.; Remington K.; Saunders R.D.C.; Scheeler F.; Shen H.;
RA Shue B.C.; Siden-Kiamos I.; Simpson M.; Skupski M.P.; Smith T.;
RA Spier E.; Spradling A.C.; Stapleton M.; Strong R.; Sun E.;
RA Svitskas R.; Tector C.; Turner R.; Venter E.; Wang A.H.; Wang X.;
RA Wang Z.-Y.; Wassarman D.A.; Weinstein G.M.; Weissbach J.;
RA Williams S.M.; Woodage T.; Worley K.C.; Wu D.; Yang S.; Yao Q.A.;
RA Ye J.; Yeh R.-F.; Zaveri J.S.; Zhan M.; Zhu G.; Zhao Q.; Zheng L.;
RA Zheng X.H.; Zhong F.N.; Zhong W.; Zhou X.; Zhu S.; Zhu X.; Smith H.O.;
RA Gibbs R.A.; Myers E.W.; Rubin G.M.; Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE-OVARY, AND IMAGINAL DISC;
RC MEDLINE=20245299; PubMed=10781936;
RA Miletich I.; Limbourg-Bouchon B.;
RT "Drosophila null slmb clones transiently deregulate Hedgehog-
independent transcription of wingless in all limb disc regeneration";
RL Mech. Dev. 93:15-26(2000).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AEO03733; AAF55853.1; -.

```

DR EMBL: AF222924; AAF63214.1; -  
 DR EMBL: AF222923; AAF63213.1; -  
 DR FlyBase: FBgn0023423; slmb.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SO SEQUENCE 510 AA; 58952 MW; F45DF126F58A012 CRC64;

Query Match 89.8%; Score 167; DB 5; Length 510;  
 Best Local Similarity 91.2%; Pred. No. 1.8e-15;  
 Matches 31; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RHCRSEKSGVYCYDQDKIVSGLRDNTIKIW 34  
 Db 202 RINCSEKSGVYCYDQDKIVSGLRDNTIKIW 235

RESULT 7  
 09BJ54 PRELIMINARY; PRT; 430 AA.  
 AC 09BJ54; 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Beta-transducin repeat-containing protein (Fragment).  
 OS Heterodera glycines (Soybean cyst nematode).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;  
 OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera;  
 OX NCB1\_TaxID=51029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kovaleva E.S., Yakovlev A.G., Masler E.P.;  
 RT "Plant parasitic nematode b-TRCP".  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF339101; AAK26376.1; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER  
 SO SEQUENCE 430 AA; 47916 MW; 4ABC3FF2DFE3A50B CRC64;

Query Match 89.2%; Score 166; DB 5; Length 430;  
 Best Local Similarity 85.3%; Pred. No. 2.2e-15;  
 Matches 29; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RHCRSEKSGVYCYDQDKIVSGLRDNTIKIW 34  
 Db 83 RINCSEKSGVYCYDQDKIVSGLRDNTIKIW 116

RESULT 8  
 0923H0 PRELIMINARY; PRT; 563 AA.  
 AC 0923H0; 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE F-box/WD40 repeat-containing protein HOS.  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N.  
 RA Bhatia N., Heter J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;  
 RT "Mouse homolog of HOS (HOS) is overexpressed in skin tumors and  
 RT implicated in constitutive activation of NF-kappaB".  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AY038079; AAK72095.1; -  
 DR MGI:2144023; FBxw1b.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR ProDom: PD000018; WD40; 4.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SO SEQUENCE 563 AA; 64741 MW; 9AB562F3FE5E3496 CRC64;

Query Match 87.6%; Score 163; DB 11; Length 563;  
 Best Local Similarity 85.3%; Pred. No. 7.6e-15;  
 Matches 29; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHCRSEKSGVYCYDQDKIVSGLRDNTIKIW 34  
 Db 253 RINCSEKSGVYCYDQDKIVSGLRDNTIKIW 286

RESULT 9  
 0922C7 PRELIMINARY; PRT; 252 AA.  
 AC 0922C7; 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Similar to f-box and WD-40 domain protein 1B (Fragment).  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: BC008552; AAO8552.1; -  
 DR MGI:2144023; FBxw1b.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 3.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER  
 SO SEQUENCE 252 AA; 28424 MW; F71737CDB7A9F75F CRC64;

Query Match 46.8%; Score 87; DB 11; Length 252;  
 Best Local Similarity 56.0%; Pred. No. 0.00022;  
 Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;



OY 10 KGVYCLQYDOKIVSGLRNTIKIM 34  
 DB 114 KGIACIQYDRDLVSGSSDNTIKIM 138

## RESULT 10

O9VZF4 PRELIMINARY: PRT: 1326 AA.

AC O9VZF4; 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE CG15010 protein (LD21332zp) (LD30271p).  
 GN AGO OR CG15010.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Plannkuch C., Baldwin D.,  
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadien E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dev I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glinka A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,  
 RA Jallali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,  
 RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Liao G., Miranda A., Mungall C.J.,  
 RA Mirando A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.,  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: COMPAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AE003480; AAC22246.1;  
 DR EMBL: AY061300; AAL28848.1;  
 DR EMBL: AY075401; AAL68231.1;  
 DR FlyBase: FBgn0041171; ago.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR000320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 2.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS0181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS0082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 1326 AA; 141360 MW; 3F42C873CFA3027F CRC64;

Query Match 40.3%; Score 75; DB 5; Length 1326;

Best Local Similarity 52.2%; Pred. No. 0.061; Mismatches 0; Gaps 0;

Matches 12; Conservative 5; Mismatches 6; Indels 0;

OY 12 VYCLQYDOKIVSGLRNTIKIM 34  
 DB 998 IYCLQYDOKIVSGLRNTIKIM 1020

## RESULT 11

IDBUN3 PRELIMINARY: PRT: 122 AA.

AC 08BUN3; 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE beta-TCP protein (Fragment).  
 DE BETA-TCP.  
 GN Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodidae; Xenopus.  
 OX NCBI\_TaxID=6335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Carnevali F.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ballarín M.;  
 RT "Analisi strutturale e funzionale del gene beta-TCP in Xenopus laevis".  
 RL Thesis (2001).  
 RL Department of Genetics and Molecular Biology "Charles Darwin",  
 RL University of Rome La Sapienza, Rome, Italy.  
 DR EMBL: AJ428938; CAD21935.1;  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 3.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 2.  
 DR SMART: SM00320; WD40; 3.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_2.  
 DR PROSITE: PS0082; WD\_REPEATS\_2; 3.  
 DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
 FT NON\_TER 1 122 122

```

SQ SEQUENCE 122 AA: 13682 MW: 36BB6B39AC8F5387 CRC64;
Query Match
Best Local Similarity 38.7%; Score 72; DB 13; Length 122;
RT 59.1%; Pred. No. 0.015;
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 10 KGVYCLQYDDQKIVSGLRDNTIK 31
:|:|||||:|||||
Db 101 KGIACLOYRDLRLVSGSSDNTI 122

RESULT 12
O96LE0 PRELIMINARY; PRT; 553 AA.
AC O96LE0;
DT 01-DEC-2001 (TREMBlrel. 15, Created)
DT 01-DEC-2001 (TREMBlrel. 15, Last sequence update)
DE CDNA FLJ11071 fis, clone PLACE1004937, moderately similar to SEL-10
protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Salto K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AK001933; BAA91986.1;
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 2.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 553 AA: 62280 MW: CA829C221986A3F2 CRC64;

Query Match
Best Local Similarity 38.7%; Score 72; DB 4; Length 553;
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIK 34
:|:|||||:|||||
Db 230 ITCLQFCGNRIYSGSDNTLTKW 252

RESULT 13
O96RI2 PRELIMINARY; PRT; 561 AA.
AC O96RI2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE F-box protein FBX30 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20003061; PubMed=10531037;
RA Winston J.T., Koepf D.M., Zhu C., Elledge S.J., Harper J.W.;
RT "A family of mammalian F-box proteins."
RL Curr. Biol. 9:1180-1182(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Koepf D.M., Winston J.T., Harper J.W., Elledge S.J.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF383178; AAK60269.1;
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRODOM: PD000018; WD40; 2.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 561 AA: 63165 MW: B81CC1E2206B0D88 CRC64;

Query Match
Best Local Similarity 38.7%; Score 72; DB 4; Length 561;
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIK 34
:|:|||||:|||||
Db 238 ITCLQFCGNRIYSGSDNTLTKW 260

RESULT 14
O96LE0 PRELIMINARY; PRT; 569 AA.
AC O96LE0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE F-box protein SEL10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Buhl A.E., Gurney M.E., Myers R.L., Shuang R., Brashler J.R., Yan R.;
RT "SEL-10 interacts with Presenilin 1, facilitates its ubiquitination,
RT and Alters A-beta production."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AY008274; AAC16640.1;
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRODOM: PD000018; WD40; 2.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 589 AA: 66120 MW: 2AFB6E8A36E68DE CRC64;

Query Match
Best Local Similarity 38.7%; Score 72; DB 4; Length 589;
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCLQYDDQKIVSGLRDNTIK 34
:|:|||||:|||||
Db 266 ITCLQFCGNRIYSGSDNTLTKW 288

```

## RESULT 15

Q96A16 PRELIMINARY; PRT; 627 AA.

AC Q96A16; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)

DE Archipelago beta form (F-box protein FBW7).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21449047; PubMed=11565033;

RA Moberg K.H., Bell D.W., Wahner D.C., Haber D.A., Hartharan I.K.;

RT "Archipelago regulates Cyclin E levels in Drosophila and is mutated in human cancer cell lines."

RL Nature 413:311-316(2001).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20003061; PubMed=10531037;

RA Winston J.T., Koepf D.M., Zhu C., Elledge S.J., Harper J.W.;

RT "A family of mammalian F-box proteins."

RL Curr. Biol. 9:1180-1182(1999).

RN [3]

RP SEQUENCE FROM N.A.

RA Koepf D.M., Winston J.T., Harper W., Elledge S.J.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

DR EMBL: AF411972; AAL06291.1; -;

DR EMBL: AY033553; AAK57547.1; -;

DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.

DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 7.

DR ProDom: PD000018; WD40; 2.

DR ProSITE: PS50181; FBOX; 1.

DR ProSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.

DR ProSITE: PS50082; WD\_REPEATS\_2; 7.

DR ProSITE: PS50294; WD\_REPEATS\_REGION; 1.

KM Repeat; WD repeat.

SQ SEQUENCE 627 AA; 70324 MW; 3D4107C053381BED CRC64;

## Query Match

Best Local Similarity 38.7%; Score 72; DB 4; Length 627;

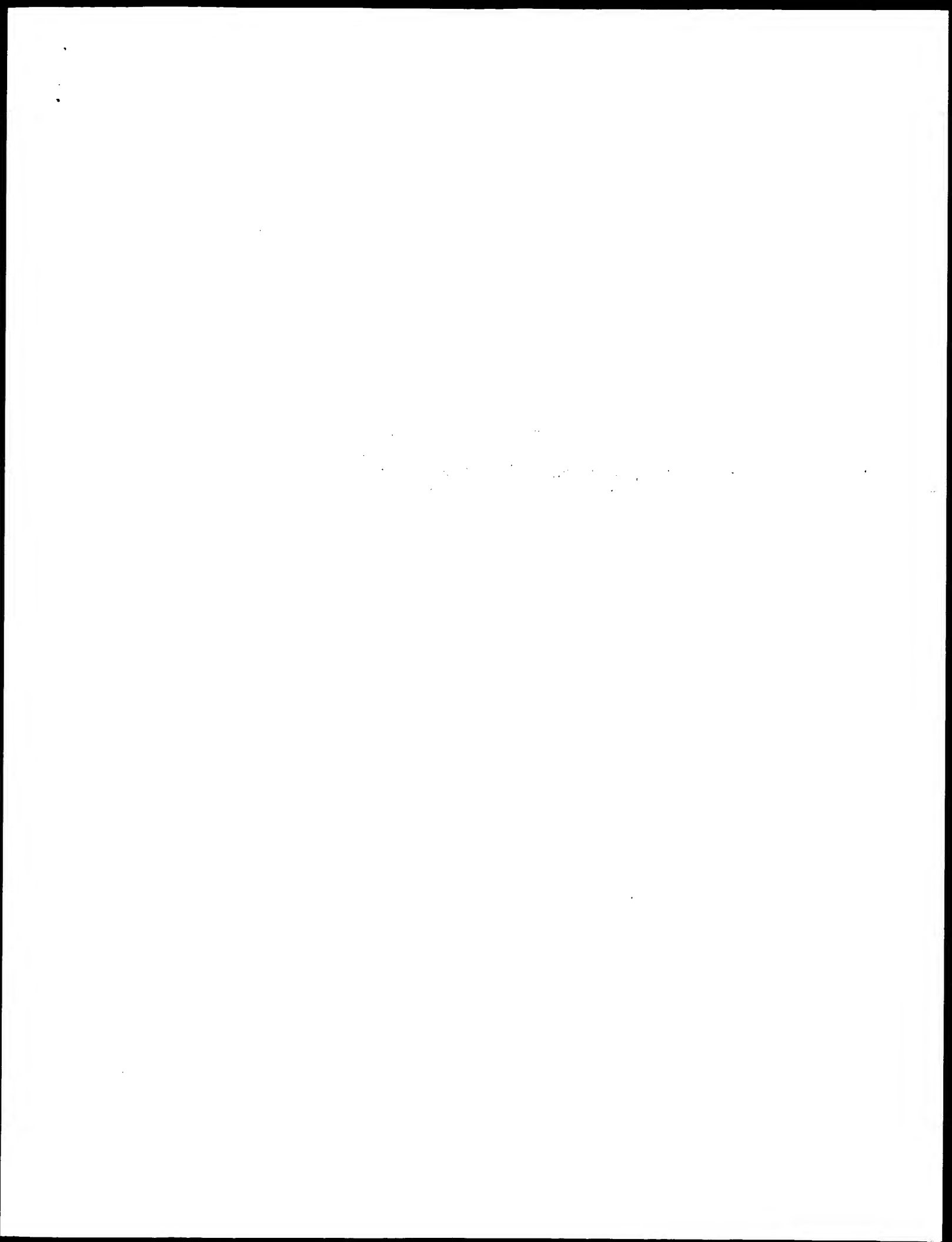
Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 12 VYCIQYDDOKIVSGLRNTIKIW 34

DB 304 ITCLQPCGNRIIVSGSDNDITLKW 326

Search completed: April 11, 2003, 11:51:32

Job time : 9.03411 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:20:44 : Search time 8.93666 Seconds  
(without alignments)  
432.406 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_304\_332

Perfect score: 153  
Sequence: 1 TGHGTSVLCIDYDERVITGSSDSTVRW 29

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A.Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	153	100.0	569	20	AAV24054
2	153	100.0	569	21	AA12812
3	153	100.0	569	21	AA12813
4	153	100.0	569	21	AAV96697
5	153	100.0	569	21	AAV83041
6	153	100.0	569	21	AAV83250
7	153	100.0	569	21	AAV83254
8	153	100.0	569	21	AAV44249
9	153	100.0	579	22	AAV78563
10	153	100.0	590	22	AAV00847

11	153	100.0	605	22	AAV78562	Human protein SEQ
12	153	100.0	608	22	AAV00960	Human bone marrow
13	153	100.0	632	22	AAV78584	Human protein SEQ
14	153	100.0	654	22	AAV79566	Human protein SEQ
15	153	100.0	654	22	AAV79567	Human protein SEQ
16	153	100.0	654	22	AAV79568	Human protein SEQ
17	152	99.3	542	21	AAV96696	Human E3 ubiquitin
18	152	99.3	542	21	AAV9127	Human protein SEQ
19	152	99.3	542	22	AAV40208	Human polypeptide
20	152	99.3	550	22	AAV41994	Human polypeptide
21	146	95.4	569	22	AAV8298	Human ZF1 protein
22	138	90.2	510	22	AAV85857	Drosophila melanog
23	133.5	87.3	28	16	AAV84936	Peptide r11 from a
24	133.5	87.3	517	16	AAV85852	WD-40 domain-contg
25	97	63.4	29	16	AAV84939	Peptide rv from a
26	86	56.2	640	21	AAV83252	F-box protein Met3
27	84	54.9	540	20	AAV22465	Human hippocampal
28	84	54.9	540	20	AAV22466	Human mammary sel-
29	84	54.9	540	22	AAV5197	Human mammary sel-
30	84	54.9	540	22	AAV59200	Human mammary sel-
31	84	54.9	545	20	AAV22464	Human hippocampal
32	84	54.9	545	20	AAV59196	Human hippocampal
33	84	54.9	553	20	AAV22463	Human hippocampal
34	84	54.9	553	22	AAV83475	Human protein sequ
35	84	54.9	553	22	AAV59195	Human mammary sel-
36	84	54.9	559	22	AAV22467	Human mammary sel-
37	84	54.9	559	22	AAV59199	Human mammary sel-
38	84	54.9	589	20	AAV22466	Human mammary sel-
39	84	54.9	589	21	AAV01204	Human GTPase assoc
40	84	54.9	589	22	AAV59198	Human mammary sel-
41	84	54.9	592	20	AAV22462	Human hippocampal
42	84	54.9	592	22	AAV59194	Human hippocampal
43	84	54.9	626	20	AAV22469	Human myc-N-sel-1
44	84	54.9	626	22	AAV59201	Protein encoded by
45	84	54.9	627	20	AAV22461	Human hippocampal

## ALIGNMENTS

RESULT 1	AAV24054	AAV24054 standard: Protein: 569 AA.
XX	AAV24054:	
AC	30-SEP-1999 (first entry)	
XX		
DE	A human beta-transducin repeat containing protein.	
XX		
XX	Beta-transducin repeat containing protein; beta-Trcp; SKP1p;	
KW	proteosome degradation pathway; Vpu protein; beta-catenin;	
KW	human immune deficiency virus-1; HIV-1; cellular protein; IkappaB;	
KW	ubiquitinylation; phosphorylated protein; tumour; apoptosis; Alzheimer's;	
KW	antiviral; antitumour; cell cycle regulation; protein degradation;	
KW	and anti-inflammatory; osteo-articular inflammation; acute inflammation;	
KW	tumour necrosis factor.	
XX		
OS	Homo sapiens.	
XX		
EH	Key	Location/Qualifiers
FT	Region	147..191
FT	Region	/note="F box sequence"
FT	Region	259..292
FT	Region	/note="WD motif"
FT	Region	304..332
FT	Region	/note="WD motif"
FT	Region	343..372
FT	Region	/note="WD motif"
FT	Region	387..415
FT	Region	/note="WD motif"
FT	Region	427..455
FT	Region	/note="WD motif"

```

FT Region 467..492
FT /note="WD motif"
FT Region 516..544
FT /note="WD motif"
PN WC9938969-A1.
XX
XX 05-AUG-1999.
XX
XX 29-JAN-1999; 99MO-FR00196.
XX
XX 09-DEC-1998; 98FR-0015545.
XX 30-JAN-1998; 98FR-0001100.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX (INSP ) INST PASTEUR.
XX
XX Arenzana Seisdedos F, Benarous R, Concordet J, Durand H;
XX Kroll M, Margoltin F;
XX WPI; 1999-469329/39.
XX N-PSDB; AAX86501.
XX
XX New human beta-transducin repeat containing protein and its
XX fragments useful as, or to screen for, antiviral, antitumour,
XX anti-inflammatory and anti-Alzheimer's agents
XX
XX Claim 1; Page 60-61; 71pp; French.
XX
XX The present sequence represents a human beta-transducin repeat containing
XX protein (beta-trcp). The protein directs proteins to the proteosome
XX degradation pathways. The protein is able to interact with the Vpu
XX protein of human immune deficiency virus-1 (HIV-1), cellular proteins
XX Ikappab or beta-catenin (bc) and/or protein Skp1. The protein controls
XX ubiquitinylation of phosphorylated proteins and thus their targeting to
XX proteosomes for degradation. Depending on whether the process is
XX inhibited or promoted, the result may be delayed breakdown of CD4 (in
XX cases of HIV-1 infection); increased activity of Ikb (and thus reduced
XX activity of NFkappab); increased degradation of mutant bc; in tumour
XX cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
XX patients. The beta-trcp protein, and its active peptide fragments, or its
XX nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
XX antitumour agents that disrupt cell cycle regulation or protein
XX degradation in human tumour cells, and anti-inflammatory agents that
XX disrupt activation by NFkappab. Fragments of the protein are also
XX useful for treating osteo-articular inflammation or acute inflammation
XX associated with release of tumour necrosis factor.
XX
XX Sequence 569 AA:
SQ
Query Match 100.0%; Score 153; DB 20; Length 569;
Best Local Similarity 100.0%; Pred. No. 1,7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGHGTGVLCLQYDERVITIGSSDSTVRW 29
DB 304 TGHGTGVLCLQYDERVITIGSSDSTVRW 332

```

```

XX
XX Mus musculus.
XX JP2000166542-A.
XX
XX 20-JUN-2000.
XX
XX 02-DEC-1998; 98JP-0343437.
XX
XX 02-DEC-1998; 98JP-0343437.
XX
XX (KAGA-) KAGAKU GIUTTSU SHINKO JIGYODAN.
XX WPI; 2000-485550/43.
XX N-PSDB; AAA73131.
XX
XX F-box protein of ubiquitin ligase SCF complex which promotes the
XX ubiquitination of Ikappab or beta-catenin
XX
XX Claim 2; Page 9-10; 19pp; Japanese.
XX
XX The present invention describes an F-box motif protein of ubiquitin
XX ligase SCF complex which promotes the ubiquitination of Ikappab or
XX beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
XX complex (SCF complex) of F-box protein containing F-box motif and WD40
XX repeat motif and has the amino acid sequence of 45 residues (AAB12811)
XX or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
XX ligase FMD1 protein) and (AAB12813, which is human beta-transducin
XX repeat containing protein (beta-trcp)). The F-box protein can be used for
XX the gene therapy of colon cancer by being recombined to a virus vector.
XX
XX Sequence 569 AA:
SQ
Query Match 100.0%; Score 153; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 1,7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGHGTGVLCLQYDERVITIGSSDSTVRW 29
DB 304 TGHGTGVLCLQYDERVITIGSSDSTVRW 332

```

```

RESULT 2
AAB12812
ID AAB12812 standard; protein: 569 AA.
XX
XX AAB12812;
XX
XX 27-NOV-2000 (first entry)
XX
XX Mouse ubiquitin ligase FMD1 protein SEQ ID NO:2.
XX
XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;
XX beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;
XX gene therapy; colon cancer; beta-transducin repeat containing protein;
XX beta-Trcp.
XX

```

```

RESULT 3
AAB12813
ID AAB12813 standard; protein: 569 AA.
XX
XX AAB12813;
XX
XX 27-NOV-2000 (first entry)
XX
XX Human beta-transducin repeat containing protein (beta-Trcp) SEQ ID NO:3.
XX
XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;
XX beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;
XX gene therapy; colon cancer; beta-transducin repeat containing protein;
XX beta-Trcp.
XX
XX Homo sapiens.
XX JP2000166542-A.
XX
XX 20-JUN-2000.
XX
XX 02-DEC-1998; 98JP-0343437.
XX
XX 02-DEC-1998; 98JP-0343437.
XX
XX (KAGA-) KAGAKU GIUTTSU SHINKO JIGYODAN.
XX WPI; 2000-485550/43.
XX N-PSDB; AAA73132.
XX
XX F-box protein of ubiquitin ligase SCF complex which promotes the
XX

```

ubiquitination of IkappaB or beta-catenin -  
XX  
XX Claim 3; Page 10-12; 19pp; Japanese.  
XX  
XX The present invention describes an F-box motif protein of ubiquitin  
CC ligase SCF complex which promotes the ubiquitination of IkappaB or  
CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
CC the gene therapy of colon cancer by being recombined to a virus vector.  
XX  
SQ Sequence 569 AA:  
Query Match 100.0%; Score 153; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TGHTGSVLCLOYDERVITIGSSDSTVRW 29  
DB 304 TGHTGSVLCLOYDERVITIGSSDSTVRW 332  
RESULT 4  
AAY96697 standard; Protein; 569 AA.  
XX  
AC AAY96697;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Human beta-TrCP.  
XX  
KW E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
KW anti-inflammatory; immunosuppressive; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200034447-A2.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99WO-US29371.  
XX  
PR 10-DEC-1998; 98US-0210060.  
XX  
PA (SIGN-) SIGNAL PHARM INC.  
PA (YISS) YISSUM RES & DEV CO.  
XX  
PI Manning AM, Mercutio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;  
PI Lavon I, Yaron A;  
XX  
DR WPI: 2000-431294/37.  
DR N-PSDB: AAA51229.  
XX  
PT Polypeptide enhancing phosphorylated IkappaB ubiquitination useful for  
PT treating disorder associated with NF-kappaB activation e.g. cancer,  
PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
PT variant  
XX  
PS Claim 21; Page 72-74; 77pp; English.  
XX  
CC Human beta-TrCP, an F-box/WD protein family member, has been shown to  
CC have homology to human E3 ubiquitin ligase (E3). E3 enhances  
CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of  
CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B  
CC degradation via the ubiquitin pathway is useful for identifying  
CC modulators of this process for use in treating diseases associated with  
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of  
CC the F-box results in a protein that functions as a dominant negative

molecule in vivo. Transient over-expression of delta-beta-TrCP (a  
CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha  
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated  
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B  
CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat  
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.  
XX  
SQ Sequence 569 AA:  
Query Match 100.0%; Score 153; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TGHTGSVLCLOYDERVITIGSSDSTVRW 29  
DB 304 TGHTGSVLCLOYDERVITIGSSDSTVRW 332  
RESULT 5  
AAY83041 standard; Protein; 569 AA.  
XX  
AC AAY83041;  
XX  
DT 16-AUG-2000 (first entry)  
XX  
DE F-box protein FBP-1.  
XX  
DE F-box protein FBP-1.  
XX  
KW F-box protein; FBP; diagnosis; treatment; screening; agonist;  
KW antagonist; proliferative disorder; differentiative disorder;  
KW breast cancer; prostate cancer; ovarian cancer; cancer;  
KW small cell lung carcinoma; immune disorder; cardiovascular disorder;  
KW inflammatory disorder; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200012679-A1.  
XX  
PD 09-MAR-2000.  
XX  
PF 27-AUG-1999; 99WO-US19560.  
XX  
PR 28-AUG-1998; 98US-0098355.  
PR 03-FEB-1999; 99US-0118568.  
PR 15-MAR-1999; 99US-0124449.  
XX  
PA (UYNY) UNIV NEW YORK STATE.  
XX  
PI Chiaur DS, Pagano M, Latres E;  
PI WPI: 2000-256635/22.  
XX  
DR N-PSDB: AA293350.  
XX  
PT Novel nucleic acid for screening compounds useful for treating  
PT proliferative and differentiative disorders such as cancer and immune  
PT disorders comprises sequences encoding ubiquitin ligases  
XX  
PS Disclosure; Figure 3a; 245pp; English.  
XX  
CC Nucleic acids encoding substrate-targeting subunits of ubiquitin  
CC ligases with F-box motifs (F-box proteins) are useful for diagnosis  
CC of proliferative and differentiative related disorders by measuring  
CC FBP gene expression. Cells expressing such proteins or  
CC their fragments are useful for screening compounds. The compounds  
CC are agonists or antagonists, which are useful for treating a  
CC proliferative or differentiative disorder in a mammal such as  
CC breast, ovarian and prostate cancer and small cell lung carcinoma  
CC and also major opportunistic infections, immune disorders,  
CC cardiovascular diseases and inflammatory disorders. FBP protein,  
CC analogs, derivatives and their subsequences, anti-FBP antibodies  
CC are also useful in diagnosis of the disorders.  
XX  
SQ Sequence 569 AA;

Query Match 100.0%; Score 153; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1,7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHGTVLCLODYDERVITITGSSDSTVRW 29  
DB 304 TGHGTVLCLODYDERVITITGSSDSTVRW 332

## RESULT 6

AAV83250  
ID AAV83250 standard; Protein; 569 AA.

AC AAV83250;

DT 16-AUG-2000 (first entry)

DE F-box protein hbeta1rcp.

XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;

KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;

KW oncoprotein; Huntington's disease; gene knockout; delivery systems;

XX human.

XX Homo sapiens.

XX MO200022110-A2.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23705.

XX 09-OCT-1998; 98US-0103787.

XX (HARD ) HARVARD COLLEGE.

XX Zhou P, Howley P;

XX WPI: 2000-317970/27.

XX N-PSDB; AA293710.

XX Targeting degradation of polypeptide useful for treating cancer and

PT other proliferative disorders, involves conjugating polypeptide with

PT ubiquitin protein ligase or inhibiting ubiquitination using organic

PT compound

XX Claim 9; Page 171; 185pp; English.

XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin

XX ligases) which can be used for the targeted degradation of a target

XX polypeptide in vivo. Targeted degradation is achieved by expressing

XX the ubiquitin ligase in a cell linked to the interaction domain of

XX the target polypeptide and thereby recruiting the target polypeptide

XX to the ubiquitin ligase. Such methods are useful for decreasing or

XX increasing the level of a target polypeptide and for creating and

XX expressing a destabilized polypeptide which is subjected to SCF

XX mediated proteolysis. Degrading any desired protein in a cell is

XX useful for preventing or treating diseases caused by the presence of

XX abnormal amount of the specific polypeptides, for drug discovery and

XX for gene therapy. Diseases treated include cancer, by degradation of

XX oncoproteins, Huntington's disease, other proliferative disorders and

XX microbial infections. The method provides a quick and easy

XX alternative to gene knockout technology. The target polypeptide can

XX be degraded at all stages, or a specific stage, of development in the

XX mature animal.

XX Sequence 569 AA;

XX Query Match 100.0%; Score 153; DB 21; Length 569;

XX Best Local Similarity 100.0%; Pred. No. 1,7e-15;

XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHGTVLCLODYDERVITITGSSDSTVRW 29  
DB 304 TGHGTVLCLODYDERVITITGSSDSTVRW 332

## RESULT 7

AAV83254  
ID AAV83254 standard; Protein; 569 AA.

AC AAV83254;

DT 16-AUG-2000 (first entry)

DE F-box protein FWD1p.

XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;

KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;

KW oncoprotein; Huntington's disease; gene knockout; delivery systems;

XX mouse; ss.

XX Mus musculus.

XX MO200022110-A2.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23705.

XX 09-OCT-1998; 98US-0103787.

XX (HARD ) HARVARD COLLEGE.

XX Zhou P, Howley P;

XX WPI: 2000-317970/27.

XX N-PSDB; AA293714.

XX Targeting degradation of polypeptide useful for treating cancer and

PT other proliferative disorders, involves conjugating polypeptide with

PT ubiquitin protein ligase or inhibiting ubiquitination using organic

PT compound

XX Claim 9; Page 184-185; 185pp; English.

XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin

XX ligases) which can be used for the targeted degradation of a target

XX polypeptide in vivo. Targeted degradation is achieved by expressing

XX the ubiquitin ligase in a cell linked to the interaction domain of

XX the target polypeptide and thereby recruiting the target polypeptide

XX to the ubiquitin ligase. Such methods are useful for decreasing or

XX increasing the level of a target polypeptide and for creating and

XX expressing a destabilized polypeptide which is subjected to SCF

XX mediated proteolysis. Degrading any desired protein in a cell is

XX abnormal amount of the specific polypeptides, for drug discovery and

XX for gene therapy. Diseases treated include cancer, by degradation of

XX oncoproteins, Huntington's disease, other proliferative disorders and

XX microbial infections. The method provides a quick and easy

XX alternative to gene knockout technology. The target polypeptide can

XX be degraded at all stages, or a specific stage, of development in the

XX mature animal.

XX Sequence 569 AA;

XX Query Match 100.0%; Score 153; DB 21; Length 569;

XX Best Local Similarity 100.0%; Pred. No. 1,7e-15;

XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



XX	13-MAY-1999;	99WMO-US10567.				
PE						
XX	13-MAY-1998;	98US-0085343.				
PR	26-AUG-1998;	98US-0098010.				
XX						
PA	(INCY- ) INCYTE PHARM INC.					
XX						
PI	Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;					
PI	Baughn MR, Yang J;					
XX						
DR	WPI: 2000-086432/07.					
DR	N-PSDB: AAZ29233.					
XX						
PT	Human cell signaling proteins useful for, e.g. diagnosing cell					
PT	proliferative and inflammatory disorders					
XX						
ES	Claim 1; Page 77-78; 90pp; English.					
XX						
CC	The present sequence is cell signalling protein-12 (CSIG-12) encoded					
CC	by cDNA obtained from Incyte clone 3239149 of COLAUCT01 library. It is					
CC	expressed in musculoskeletal, gastrointestinal and nervous tissues and is					
CC	found to be homologous to beta-transducin repeats containing					
CC	hybridisation probe for detecting CSIG related sequences or allelic					
CC	variants. Recombinant CSIG can be produced in host cells by transforming					
CC	them with genetically engineered vectors. Agonists or antagonists can be					
CC	used in the treatment of cell proliferative and inflammatory disorders					
CC	associated with decreased or increased CSIG expression. CSIG is used in					
CC	the diagnosis, prevention and treatment of cell proliferative disorders					
CC	like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory					
CC	disorders like AIDS, Addison's disease, multiple sclerosis, etc.					
XX						
SQ	Sequence	569 AA;				
	Query Match	100.0%; Score 153; DB 21;				
	Best Local Similarity	100.0%; Pred. No. 1.7e-15;				
	Matches	29; Conservative	0; Mismatches	0; Indels	0; Gaps	0
OY	1 TGGTGSVLCLOYDERVITIGSSDSIVRW	29				
DB	304 TGGTGSVLCLOYDERVITIGSSDSIVRW	332				
	RESULT 9					
	AAW78583					
XX	AAW78583 standard; Protein; 579 AA.					
XX						
AC	AAW78583;					
XX						
DT	06-NOV-2001 (first entry)					
XX						
XX	Human protein SEQ ID NO 1245.					
XX						
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;					
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;					
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;					
KW	nervous system disorder; arthritis; inflammation.					
OS						
OS	Homo sapiens.					
XX						
PN	WO200157190-A2.					
XX						
PD	09-AUG-2001.					
XX						
PF	05-FEB-2001; 2001WO-US04098.					
XX						
PR	03-FEB-2000; 2000US-0496914.					
PR	27-JUN-2000; 2000US-0560875.					
PR	20-JUN-2000; 2000US-0598075.					
PR	19-JUL-2000; 2000US-0620325.					
PR	01-SEP-2000; 2000US-0654936.					
PR	15-SEP-2000; 2000US-0663561.					

PR 20-OCT-2000; 2000US-0693325.  
PR 30-NOV-2000; 2000US-0728422.  
XX  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Gao Y, Ma Y;  
PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R,  
XX  
XX WPI; 2001-476283/51.  
DR N-PSDB; AAK51716.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
XX  
PS Claim 20; Page 3504-3505; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53425) and the  
CC encoded polypeptides (AAM78323-AAK60302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAM60020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX  
XX  
SQ Sequence 579 AA;  
XX  
XX  
Query Match 100.0%; Score 153; DB 22; Length 579;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGHGTGVLQYDERVITITGSSDSTVRVW 29  
DB 314 TGHGTGVLQYDERVITITGSSDSTVRVW 342  
XX  
XX  
RESULT 10  
AAM00847  
ID AAM00847 standard; Protein; 590 AA.  
XX  
XX  
AC AAM00847;  
XX  
XX  
DT 01-OCT-2001 (first entry)  
XX  
XX Human bone marrow protein, SEQ ID NO: 210.  
DE  
XX  
XX Human; bone marrow; antiinflammatory; cytosolic; neuroprotective;  
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;  
KW immunosuppressive; gene therapy; cytokine cell proliferation;  
KW cell differentiation modulator; immune disorder; infection; cancer;  
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200153453-A2.  
PN  
XX  
XX 26-JUL-2001.  
PD  
XX  
XX 23-DEC-2000; 2000WO-US34960.  
PE  
XX  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.  
PR 30-NOV-2000; 2000US-0250583.  
XX  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX  
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
PI Ren F, Wang J, Wejhrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Drmanac RT;  
XX  
XX WPI; 2001-488707/53.  
DR N-PSDB; AAH89966.  
XX  
XX  
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
PT for treating e.g. cancer and immune deficiency disorders -  
XX  
XX  
PS Claim 10; Page 354-355; 648pp; English.  
XX  
XX  
XX The present sequence is one of 251 novel human polypeptides encoded  
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the  
CC polypeptide encoded by it are useful in the treatment of various  
CC immune deficiencies and disorders. The deficiencies and disorders may  
CC be genetic, or may result from a viral (e.g. HIV), bacterial or fungal  
CC infection, or may result from an autoimmune disorder, a coagulation  
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
CC suppression of an inflammatory response or treatment of a nervous  
CC system disorder such as Alzheimer's disease. Detection of the presence  
CC or increased expression of the polynucleotide or the protein it  
CC encodes is useful for the diagnosis and/or prognosis of one  
CC or more types of cancer. The polynucleotide and polypeptide can be  
CC used as nutritional sources or supplements and in the screening of  
CC chemical compounds as potential drugs.  
XX  
XX  
SQ Sequence 590 AA;  
XX  
XX  
Query Match 100.0%; Score 153; DB 22; Length 590;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGHGTGVLQYDERVITITGSSDSTVRVW 29  
DB 325 TGHGTGVLQYDERVITITGSSDSTVRVW 353  
XX  
XX  
RESULT 11  
AAM78582  
ID AAM78582 standard; Protein; 605 AA.  
XX  
XX  
AC AAM78582;  
XX  
XX  
DT 06-NOV-2001 (first entry)  
XX  
XX  
XX Human protein SEQ ID NO 1244.  
DE  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157190-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 05-FEB-2001; 2001WO-US04098.  
PE  
XX  
XX 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0654936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Dirmnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
XX WPI: 2001-476283/51.  
DR N-PSDB; AAK51715.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
XX Claim 20; Page 3503-3504; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAW80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX  
XX SQ Sequence 605 AA;  
XX  
XX Query Match 100.0%; Score 153; DB 22; Length 605;  
XX Best Local Similarity 100.0%; Pred. No. 1.8e-15;  
XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TGHGTGVLCLQYDERVITITGSSDSTVRW 29  
DB 340 TGHGTGVLCLQYDERVITITGSSDSTVRW 368  
RESULT 12  
AAW00960  
ID AAW00960 standard; Protein; 608 AA.  
XX  
XX AAW00960;  
XX  
XX 01-OCT-2001 (first entry)  
XX  
XX Human bone marrow protein, SEQ ID NO: 436.  
XX  
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;  
XX antifungal; antibacterial; antifungal; anti-HIV; haemostatic;  
XX immunosuppressive; gene therapy; cytokine cell proliferation;  
XX cell differentiation modulator; immune disorder; infection; cancer;  
XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.  
XX  
XX Homo sapiens.  
XX  
XX WO200153453-A2.  
XX  
XX 26-JUL-2001.  
XX  
XX 23-DEC-2000; 2000WO-US34960.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX 09-JUL-2000; 2000US-0598042.  
XX 19-JUL-2000; 2000US-0620312.  
XX 03-AUG-2000; 2000US-0653450.  
XX 14-SEP-2000; 2000US-0662191.  
XX 19-OCT-2000; 2000US-0693036.

PR 30-NOV-2000; 2000US-0250583.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
PI Ren F, Wang J, Wejhrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Dirmnac RT;  
XX  
XX WPI: 2001-488707/53.  
DR N-PSDB; AAH90079.  
XX  
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
PT for treating e.g. cancer and immune deficiency disorders -  
XX  
XX Claim 10; Page 523-524; 648pp; English.  
XX  
XX The present sequence is one of 251 novel human polypeptides encoded  
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the  
CC polypeptide encoded by it are useful in the treatment of various  
CC immune deficiencies and disorders. The deficiencies and disorders may  
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal  
CC infection, or may result from an autoimmune disorder, a coagulation  
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
CC suppression of an inflammatory response or treatment of a nervous  
CC system disorder such as Alzheimer's disease. Detection of the presence  
CC or increased expression of the polynucleotide or the protein it  
CC encodes is useful for the diagnosis and/or prognosis of one  
CC or more types of cancer. The polynucleotide and polypeptide can be  
CC used as nutritional sources or supplements and in the screening of  
CC chemical compounds as potential drugs.  
XX  
XX SQ Sequence 608 AA;  
XX  
XX Query Match 100.0%; Score 153; DB 22; Length 608;  
XX Best Local Similarity 100.0%; Pred. No. 1.8e-15;  
XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TGHGTGVLCLQYDERVITITGSSDSTVRW 29  
DB 343 TGHGTGVLCLQYDERVITITGSSDSTVRW 371  
RESULT 13  
AAW78584  
ID AAW78584 standard; Protein; 632 AA.  
XX  
XX AAW78584;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
XX Human protein SEQ ID NO 1246.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX nervous system disorder; arthritis; inflammation.  
XX  
XX Homo sapiens.  
XX  
XX WO200157190-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 05-FEB-2001; 2001WO-US04098.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
XX 27-APR-2000; 2000US-0560875.  
XX 20-JUN-2000; 2000US-0598075.  
XX 19-JUL-2000; 2000US-0620325.  
XX 01-SEP-2000; 2000US-0654936.  
XX 15-SEP-2000; 2000US-0663561.  
XX 20-OCT-2000; 2000US-0693325.  
XX 30-NOV-2000; 2000US-0728422.

DR WPI; 2001-476283/51.  
DR N-PSDB; AAK51717.

Claim 20; Page 3505-3507; 6221pp; English.

AA	Sequence	632 AA;
SQ		

## RESULT 14

AAC  
YY  
AAM79566;

Human protein SEQ ID NO 3212.

Homo sapiens.

WO200157190-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US04098.

03-FEB-2000; 2000US-0496914  
27-APR-2000; 2000US-0560875  
20-JUN-2000; 2000US-0598075  
19-JUL-2000; 2000US-0620325  
01-SEP-2000; 2000US-0654936  
15-SEP-2000; 2000US-0665361  
20-OCT-2000; 2000US-0693325  
30-NOV-2000; 2000US-0728422.

(HYSE-) HYSEQ INC.

WPI; 2001-476283/51.  
N-PSDB; AAK52699.

Claim 20; Page 285-286; 6221pp; English.

Sequence 654 AA;

## RESULT 15

AAM79567;

Human protein SEQ ID NO 3213.

Homo sapiens.

MO200157190-A2

09-AUG-2001.

05-FEB-2001; 2001WO-US04098.

03-FEB-2000; 2000US-0406914  
27-APR-2000; 2000US-0560875  
20-JUN-2000; 2000US-0538075  
19-JUL-2000; 2000US-0620325  
15-SEP-2000; 2000US-0654936  
11-SEP-2000; 2000US-0663561  
20-OCT-2000; 2000US-0693325  
30-NOV-2000; 2000US-0728422

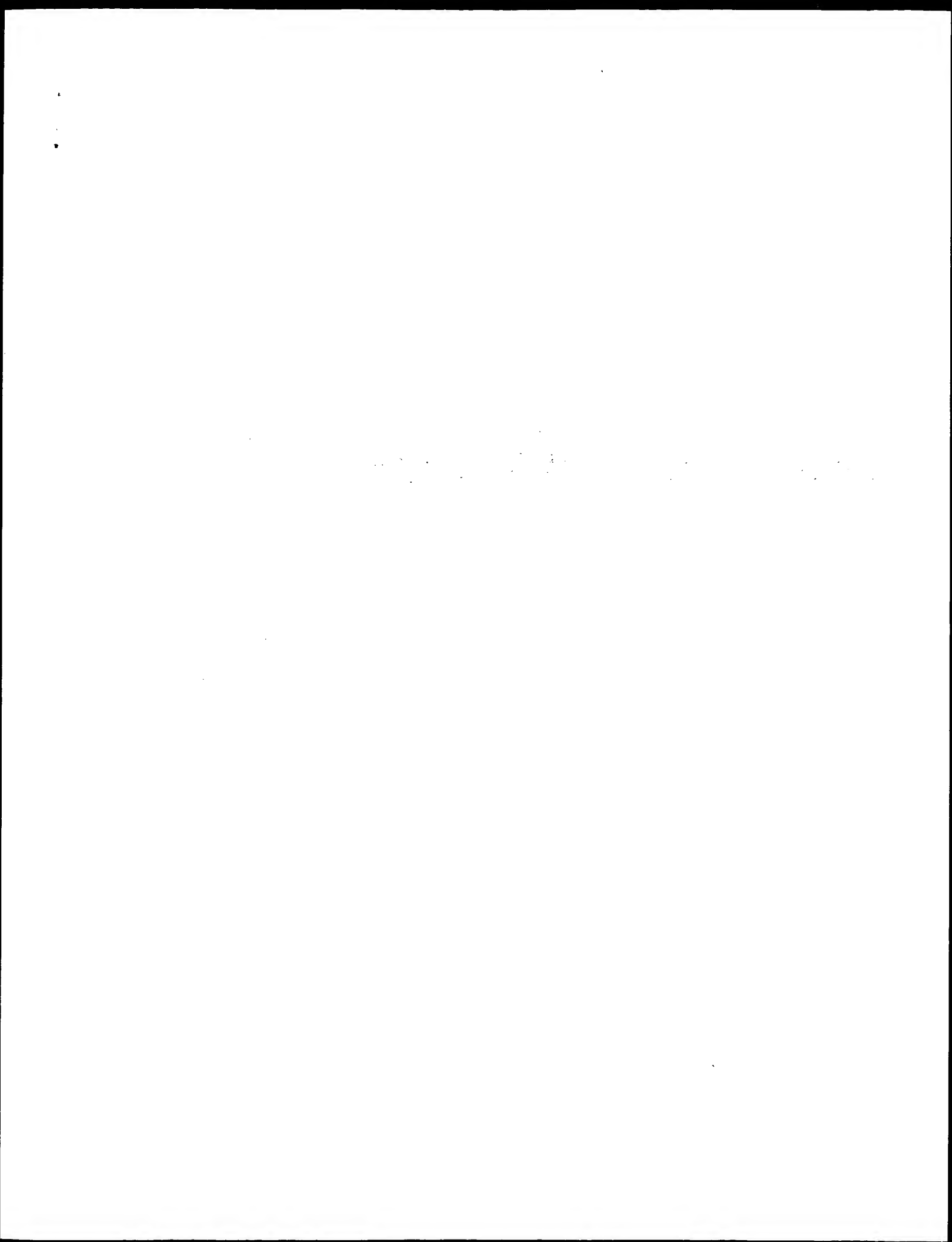
(HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmenac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI; 2001-476283/51.  
 DR N-PSDB; AAK52700.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20; Page 286; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX  
 SO Sequence 654 AA;

Query Match 100.0%; Score 153; DB 22; Length 654;  
 Best Local Similarity 100.0%; Pred. No. 2e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHGTVLCLOYDERVITIGSSDSTVRW 29  
 |||||||||||||||||||||||||||||  
 Db 389 TGHGTVLCLOYDERVITIGSSDSTVRW 417

Search completed: April 11, 2003, 11:48:14  
 Job time : 9.93666 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 3.00244 Seconds  
(without alignments)  
284.191 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_304\_332

Perfect score: 153  
Sequence: 1 TGHGTVLCLOYDERVITGSSDSTVRW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574.

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCRNUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	133.5	87.3	28	1	US-08-190-802A-84
2	133.5	87.3	28	4	US-08-477-346-84
3	133.5	87.3	28	4	US-08-473-089-84
4	133.5	87.3	28	4	US-08-487-072A-84
5	133.5	87.3	517	1	US-08-190-802A-30
6	133.5	87.3	517	4	US-08-477-346-30
7	133.5	87.3	517	4	US-08-473-089-30
8	133.5	87.3	517	4	US-08-487-072A-30
9	97	63.4	29	1	US-08-190-802A-87
10	97	63.4	29	4	US-08-477-346-87
11	97	63.4	29	4	US-08-473-089-87
12	97	63.4	29	4	US-08-487-072A-87
13	86	56.2	640	4	US-09-177-165A-30
14	82	53.6	209	3	US-08-899-578-6
15	81	52.9	587	3	US-08-899-578-2
16	78	51.0	325	2	US-08-828-922-3
17	77	50.3	34	1	US-08-190-802A-83
18	77	50.3	34	4	US-08-477-346-83
19	77	50.3	34	4	US-08-473-089-83
20	77	50.3	34	4	US-08-487-072A-83
21	75	49.0	31	1	US-08-190-802A-249
22	75	49.0	31	4	US-08-477-346-249
23	75	49.0	31	4	US-08-473-089-249
24	75	49.0	31	4	US-08-487-072A-249
25	75	49.0	514	1	US-08-190-802A-66
26	75	49.0	514	4	US-08-477-346-66
27	75	49.0	514	4	US-08-473-089-66

28	75	49.0	514	4	US-08-487-072A-66	Sequence 66, App1
29	73.5	48.0	41	4	US-08-988-856B-23	Sequence 23, App1
30	73.5	48.0	779	1	US-08-190-802A-32	Sequence 32, App1
31	73.5	48.0	779	4	US-08-477-346-32	Sequence 32, App1
32	73.5	48.0	779	4	US-08-473-089-32	Sequence 32, App1
33	73.5	48.0	779	4	US-08-487-072A-32	Sequence 32, App1
34	73.5	48.0	779	4	US-09-177-165A-29	Sequence 29, App1
35	73	47.7	31	1	US-08-190-802A-93	Sequence 93, App1
36	73	47.7	31	1	US-08-190-802A-196	Sequence 196, App1
37	73	47.7	31	1	US-08-477-346-93	Sequence 93, App1
38	73	47.7	31	4	US-08-477-346-196	Sequence 196, App1
39	73	47.7	31	4	US-08-473-089-93	Sequence 93, App1
40	73	47.7	31	4	US-08-473-089-196	Sequence 196, App1
41	73	47.7	31	4	US-08-487-072A-93	Sequence 93, App1
42	73	47.7	31	4	US-08-487-072A-196	Sequence 196, App1
43	73	47.7	422	1	US-08-190-802A-52	Sequence 52, App1
44	73	47.7	422	4	US-08-477-346-52	Sequence 52, App1
45	73	47.7	422	4	US-08-473-089-52	Sequence 52, App1

## ALIGNMENTS

## RESULT 1

US-08-190-802A-84

Sequence 84, Application US/08190802A

Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

TITLE OF INVENTION: WO-40 - Derived peptides and Uses

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190, 802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 84:

SEQUENCE CHARACTERISTICS:

LENGTH: 28 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP I1, Fig. 13

US-08-190-802A-84

Query Match 87.3%; Score 133.5; DB 1; Length 28;  
Best Local Similarity 96.4%; Pred. No. 2.7e-13;  
Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

2 TGHGTVLCLOYDERVITGSSDSTVRW 29

Db 1 GHTGSVLCQYDERVITIG-SDSTVRW 27

## RESULT 2

US-08-477-346-84  
; Sequence 84, Application US/08477346  
; Patent No. 6262023  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,346  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,072  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; INFORMATION FOR SEQ ID NO: 84:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: BETA TRCP rII, Fig. 13  
US-08-477-346-84  
Query Match 87.3%; Score 133.5; DB 4; Length 28;  
Best Local Similarity 96.4%; Pred. No. 2.7e-13;  
Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Db 1 GHTGSVLCQYDERVITIG-SDSTVRW 27

## RESULT 3

US-08-473-089-84  
; Sequence 84, Application US/08473089  
; Patent No. 6342368  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rII, Fig. 13  
US-08-473-089-84

Query Match 87.3%; Score 133.5; DB 4; Length 28;  
Best Local Similarity 96.4%; Pred. No. 2.7e-13;  
Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Db 1 GHTGSVLCQYDERVITIG-SDSTVRW 27

## RESULT 4

US-08-487-072A-84  
; Sequence 84, Application US/08487072A  
; Patent No. 6423684  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,072A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:



TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 INFORMATION FOR SEQ ID NO: 84:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 28 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: BETA TRCP r11, Fig. 13  
 US-08-487-072A-84

Query Match 87.3%; Score 133.5; DB 4; Length 28;  
 Best Local Similarity 96.4%; Pred. No. 2,7e-13;  
 Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy 2 GHTGSVLCLOYDERVITIGSSDSSTVRW 29  
 Db 1 GHTGSVLCLOYDERVITIG-SDSTVRW 27

RESULT 5  
 US-08-190-802A-30  
 Sequence 30, Application US/08190802A  
 Patent No. 3519003  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 APPLICANT: Ron, Dorit  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thereof  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dehlinger & Associates  
 STREET: P.O. Box 60850  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94306-0850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/190, 802A  
 FILING DATE: 01-FEB-1994  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fabian, Gary R.  
 REGISTRATION NUMBER: 33,875  
 REFERENCE/DOCKET NUMBER: 8600-0139  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 30:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 517 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
 US-08-190-802A-30

Query Match 87.3%; Score 133.5; DB 1; Length 517;  
 Best Local Similarity 96.4%; Pred. No. 7.5e-12;  
 Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 Oy 2 GHTGSVLCLOYDERVITIGSSDSSTVRW 29

Db 270 GHTGSVLCLOYDERVITIG-SDSTVRW 296

RESULT 6  
 US-08-477-346-30  
 Sequence 30, Application US/08477346  
 Patent No. 6262023  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 APPLICANT: Ron, Dorit  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thereof  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 2000 Pennsylvania Avenue, NW  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20006-1812  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/477,346  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/487,072  
 FILING DATE: 07-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MURASHIGE, KATE H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 2550-0025.20  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 INFORMATION FOR SEQ ID NO: 30:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 517 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
 US-08-477-346-30

Query Match 87.3%; Score 133.5; DB 4; Length 517;  
 Best Local Similarity 96.4%; Pred. No. 7.5e-12;  
 Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy 2 GHTGSVLCLOYDERVITIGSSDSSTVRW 29  
 Db 270 GHTGSVLCLOYDERVITIG-SDSTVRW 296

RESULT 7  
 US-08-473-089-30  
 Sequence 30, Application US/08473089  
 Patent No. 6342368  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 APPLICANT: Ron, Dorit  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thereof  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-473-089-30

Query Match 87.3%; Score 133.5; DB 4; Length 517;  
Best Local Similarity 96.4%; Pred. No. 7.5e-12;  
Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 GHGGSVLCLOYDERVITGSSDSSTVRW 29  
Db 270 GHGGSVLCLOYDERVITG-SDSTVRW 296

RESULT 8  
US-08-487-072A-30  
Sequence 30, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-487-072A-30

Query Match 87.3%; Score 133.5; DB 4; Length 517;  
Best Local Similarity 96.4%; Pred. No. 7.5e-12;  
Matches 27; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 GHGGSVLCLOYDERVITGSSDSSTVRW 29  
Db 270 GHGGSVLCLOYDERVITG-SDSTVRW 296

RESULT 9  
US-08-190-802A-87  
Sequence 87, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Denlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fadian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rV, Fig. 13  
US-08-190-802A-87

Query Match 63.4%; Score 97; DB 1; Length 29;  
Best Local Similarity 50.0%; Pred. No. 7e-08;  
Matches 14; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 2 GHGGSVLCLOYDERVITGSSDSSTVRW 29

Db 1 GHRGICACLOYDRRLVVGSSDNTIRLW 28

RESULT 10  
US-08-477-346-87  
Sequence 87, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theeof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-0763  
TELEFAX: (202) 887-1500  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13  
US-08-477-346-87  
Query Match 63.4%; Score 97; DB 4; Length 29;  
Best Local Similarity 50.0%; Pred. No. 7e-08;  
Matches 14; Conservative 9; Mismatches 5; Indels 0; Gaps 0;  
QY 2 GHTGSVLCLOYDERVITIGSSDSTVRW 29  
DB 1 GHRGICACLOYDRRLVVGSSDNTIRLW 28  
RESULT 11  
US-08-473-089-87  
Sequence 87, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theeof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-0763  
TELEFAX: (202) 887-1500  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13  
US-08-473-089-87  
Query Match 63.4%; Score 97; DB 4; Length 29;  
Best Local Similarity 50.0%; Pred. No. 7e-08;  
Matches 14; Conservative 9; Mismatches 5; Indels 0; Gaps 0;  
QY 2 GHTGSVLCLOYDERVITIGSSDSTVRW 29  
DB 1 GHRGICACLOYDRRLVVGSSDNTIRLW 28  
RESULT 12  
US-08-487-072A-87  
Sequence 87, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theeof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

MOLECULE TYPE: protein  
US-08-899-578-2

Fri Apr 11 13:31:09 2003

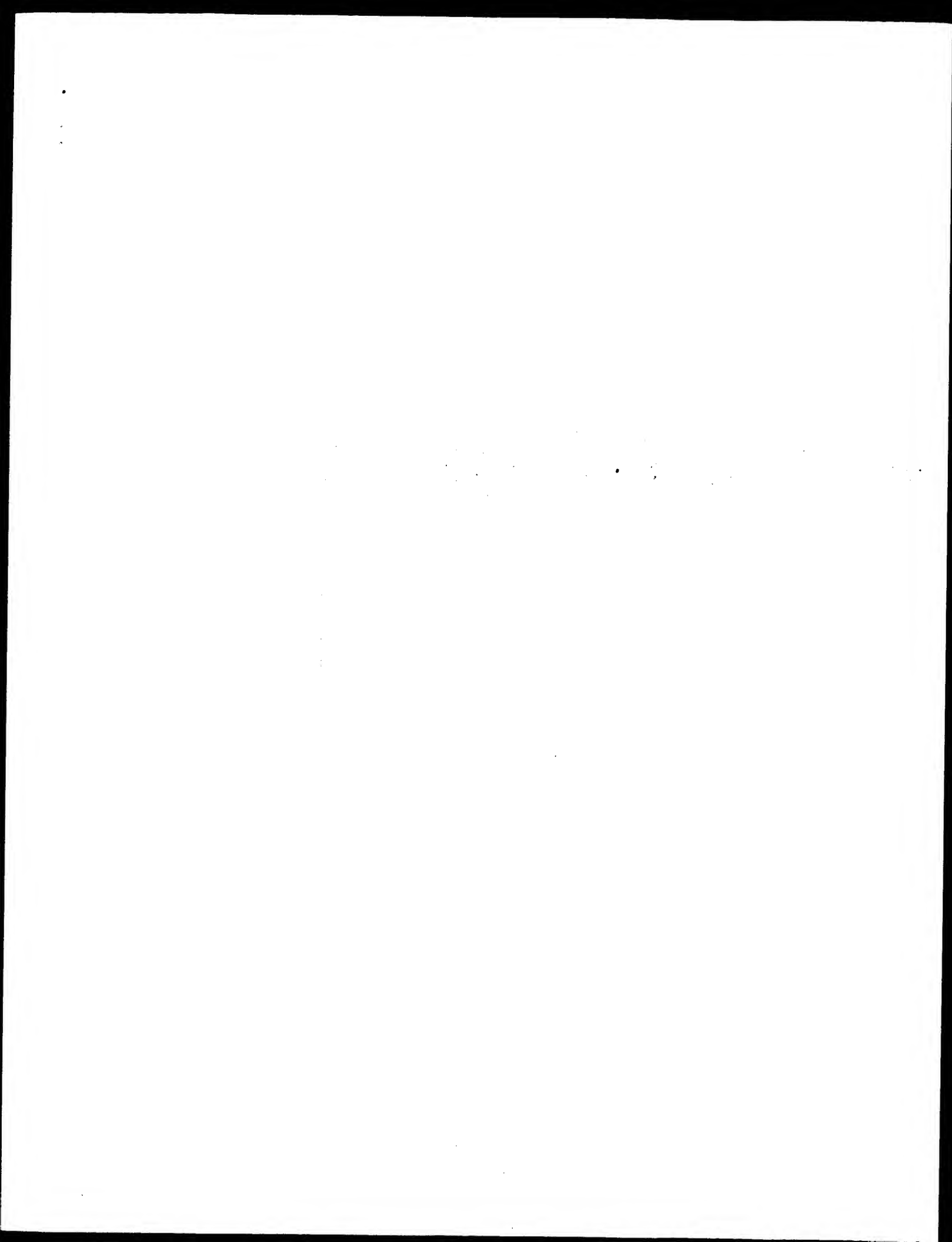
us-09-601-168b-2\_copy\_304\_332.ra

Page 7

Query Match 52.9%; Score 81; DB 3; Length 587;  
Best Local Similarity 46.4%; Pred. No. 0.00049;  
Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHTGSVLCLOYDERVITIGSSDSSTVRW 29  
||| : | :  
Db 336 GHTSTVRNCMAMAGSILVTGSRDTTLRW 363

Search completed: April 10, 2003, 12:51:18  
Job time : 4.00244 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:49:06 ; Search time 2.96711 Seconds  
(without alignments)  
597.532 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_304\_332  
Perfect score: 153  
Sequence: 1 TGHGTVLCIQDYDERVITGSSDSIVRW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	100.0	569	9 US-10-038-010-8	Sequence 8, Appl1
2	153	100.0	569	12 US-10-042-417-2	Sequence 2, Appl1
3	98	64.1	678	10 US-09-801-368-314	Sequence 30, Appl1
4	86	56.2	640	9 US-10-060-019-30	Sequence 3, Appl1
5	84	54.9	540	9 US-09-213-888-7	Sequence 7, Appl1
6	84	54.9	540	9 US-09-213-888-10	Sequence 10, Appl1
7	84	54.9	540	9 US-09-328-877A-7	Sequence 7, Appl1
8	84	54.9	540	9 US-09-328-877A-10	Sequence 10, Appl1
9	84	54.9	545	9 US-09-213-888-6	Sequence 6, Appl1
10	84	54.9	545	9 US-09-328-877A-6	Sequence 6, Appl1
11	84	54.9	553	9 US-09-213-888-5	Sequence 5, Appl1
12	84	54.9	553	9 US-09-328-877A-5	Sequence 5, Appl1
13	84	54.9	559	9 US-09-213-888-9	Sequence 9, Appl1
14	84	54.9	559	9 US-09-328-877A-9	Sequence 9, Appl1
15	84	54.9	589	9 US-09-213-888-8	Sequence 8, Appl1
16	84	54.9	589	9 US-09-328-877A-8	Sequence 8, Appl1
17	84	54.9	592	9 US-09-213-888-4	Sequence 4, Appl1
18	84	54.9	592	9 US-09-328-877A-4	Sequence 4, Appl1
19	84	54.9	626	9 US-09-213-888-21	Sequence 21, Appl1

20	84	54.9	626	9 US-09-328-877A-21	Sequence 21, Appl1
21	84	54.9	627	9 US-09-213-888-3	Sequence 3, Appl1
22	84	54.9	627	9 US-09-328-877A-3	Sequence 3, Appl1
23	84	54.9	666	9 US-09-213-888-27	Sequence 27, Appl1
24	84	54.9	666	9 US-09-328-877A-27	Sequence 27, Appl1
25	84	54.9	669	9 US-09-213-888-25	Sequence 25, Appl1
26	84	54.9	669	9 US-09-328-877A-25	Sequence 25, Appl1
27	79	51.6	422	12 US-10-042-417-4	Sequence 4, Appl1
28	73.5	48.0	41	10 US-09-843-845-23	Sequence 23, Appl1
29	73.5	48.0	9	9 US-10-083-357-897	Sequence 897, Appl1
30	73.5	48.0	9	9 US-10-060-019-29	Sequence 29, Appl1
31	73	47.7	744	10 US-09-925-300-1347	Sequence 1347, Appl1
32	72.5	47.4	64	10 US-09-843-845-20	Sequence 20, Appl1
33	72	47.1	261	9 US-10-132-744A-4	Sequence 4, Appl1
34	72	47.1	484	9 US-10-132-744A-2	Sequence 2, Appl1
35	72	47.1	485	9 US-10-132-744A-6	Sequence 6, Appl1
36	72	47.1	1146	9 US-09-832-292-10	Sequence 10, Appl1
37	72	47.1	1146	10 US-09-994-485-6	Sequence 6, Appl1
38	70.5	46.1	425	9 US-09-893-519A-25	Sequence 25, Appl1
39	70.5	46.1	466	10 US-09-925-301-997	Sequence 997, Appl1
40	70	45.8	67	10 US-09-864-761-46849	Sequence 46849, A
41	70	45.8	114	9 US-09-796-692-1445	Sequence 1445, Ap
42	70	45.8	114	9 US-09-796-692-1541	Sequence 1541, Ap
43	70	45.8	114	9 US-09-796-692-1895	Sequence 1895, Ap
44	70	45.8	316	10 US-09-828-310-12	Sequence 12, Appl1
45	70	45.8	317	9 US-09-998-042-9	Sequence 9, Appl1

## ALIGNMENTS

RESULT 1  
US-10-038-010-8 Application US/10038010  
Sequence 8, Appl1  
Publication No. US20030040089A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS  
APPLICANT: PIERRE, Legrain  
TITLE OF INVENTION: Protein-protein interactions in adipocyte cells  
FILE REFERENCE: B4767A  
CURRENT APPLICATION NUMBER: US/10/038,010  
CURRENT FILING DATE: 2002-07-23  
PRIOR APPLICATION NUMBER: US 60/259,377  
PRIOR FILING DATE: 2001-01-02  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: beta-Trip1  
LOCATION: (1)..(569)  
OTHER INFORMATION:  
US-10-038-010-8

Query Match 100.0% Score 153; DB 9; Length 569;  
Best Local Similarity 100.0%; Pred. NO. 1e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHGTVLCIQDYDERVITGSSDSIVRW 29  
Db 304 TGHGTVLCIQDYDERVITGSSDSIVRW 332

RESULT 2  
US-10-042-417-2  
Sequence 2, Appl1 Application US/10042417  
Patent No. US20020123082A1  
GENERAL INFORMATION:  
APPLICANT: Pagano, M.  
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS

FILE REFERENCE: 5914-090-999  
CURRENT APPLICATION NUMBER: US/10/042,417  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-5  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-042-417-2

Query Match 100.0%; Score 153; DB 12; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHTGSVLCLODYDERVITIGSSDSYRW 29  
DB 304 TGHTGSVLCLODYDERVITIGSSDSYRW 332

RESULT 3  
US-09-801-368-314  
Sequence 314, Application US/09801368  
Patent No. US20020128250A1  
GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Call, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 314  
LENGTH: 678  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
US-09-801-368-314

Query Match 64.1%; Score 98; DB 10; Length 678;  
Best Local Similarity 46.4%; Pred. No. 4.8e-07;  
Matches 13; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 2 GHTGSVLCLODYDERVITIGSSDSYRW 29  
DB 347 GHTGSVLCLODYDERVITIGSSDSYRW 374

RESULT 4  
US-10-060-019-30  
Sequence 30, Application US/10060019  
Publication No. US2003003564A1  
GENERAL INFORMATION:  
APPLICANT: Tyers, Mike  
APPLICANT: Williams, Andrew  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN

TITLE OF INVENTION: DEPENDENT PROTEOLYSIS  
FILE REFERENCE: 11757.100S01  
CURRENT APPLICATION NUMBER: US/10/060,019  
CURRENT FILING DATE: 2002-01-29  
PRIOR APPLICATION NUMBER: US/09/177,165  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: 60/092,443  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/063,254  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 30  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-060-019-30

Query Match 56.2%; Score 86; DB 9; Length 640;  
Best Local Similarity 51.7%; Pred. No. 3.3e-05;  
Matches 15; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 TGHTGSVLCLODYDERVITIGSSDSYRW 29  
DB 339 SGHSDGVKTLVFDKRLITGSLDKTIRW 367

RESULT 5  
US-09-213-888-7  
Sequence 7, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-7

Query Match 54.9%; Score 84; DB 9; Length 540;  
Best Local Similarity 46.4%; Pred. No. 5.6e-05;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHTGSVLCLODYDERVITIGSSDSYRW 29  
DB 292 GHTSTVRCHMLHKKRVSGSRDATLRW 319

RESULT 6  
US-09-213-888-10  
Sequence 10, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27



SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 540  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-213-888-10

Query Match 54.9%; Score 84; DB 9; Length 540;  
Best Local Similarity 46.4%; Pred. No. 5.6e-05;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 2 GHTGSVLCLOYDERVITITGSSDSTVRYW 29  
Db 292 GHTSTVRCMHLEKRVVSGSRDATTLRW 319

RESULT 7  
US-09-328-877A-7

; Sequence 7, Application US/09328877A  
; Patent No. US20020177187A1  
; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Li, Jinhe

; APPLICANT: Pauley, Adele M.

; APPLICANT: Pharmacia & Upjohn Company

; TITLE OF INVENTION: Encode Them

; FILE REFERENCE: 6142

; CURRENT APPLICATION NUMBER: US/09/328, 877A

; CURRENT FILING DATE: 1999-06-09

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 7

; LENGTH: 540

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-328-877A-7

Query Match 54.9%; Score 84; DB 9; Length 540;  
Best Local Similarity 46.4%; Pred. No. 5.6e-05;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 2 GHTGSVLCLOYDERVITITGSSDSTVRYW 29  
Db 292 GHTSTVRCMHLEKRVVSGSRDATTLRW 319

RESULT 8  
US-09-328-877A-10

; Sequence 10, Application US/09328877A  
; Patent No. US20020177187A1  
; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Li, Jinhe

; APPLICANT: Pauley, Adele M.

; APPLICANT: Pharmacia & Upjohn Company

; TITLE OF INVENTION: Encode Them

; FILE REFERENCE: 6142

; CURRENT APPLICATION NUMBER: US/09/328, 877A

; CURRENT FILING DATE: 1999-06-09

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 10

; LENGTH: 540

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-328-877A-10

Query Match 54.9%; Score 84; DB 9; Length 540;  
Best Local Similarity 46.4%; Pred. No. 5.6e-05;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 2 GHTGSVLCLOYDERVITITGSSDSTVRYW 29  
Db 292 GHTSTVRCMHLEKRVVSGSRDATTLRW 319

RESULT 9  
US-09-213-888-6

; Sequence 6, Application US/09213888A  
; Patent No. US20020164683A1  
; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Li, Jinhe

; APPLICANT: Pauley, Adele M.

; APPLICANT: Pharmacia & Upjohn Company

; TITLE OF INVENTION: Encode Them

; FILE REFERENCE: 6142

; CURRENT APPLICATION NUMBER: US/09/213, 888A

; CURRENT FILING DATE: 1998-12-17

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 6

; LENGTH: 545

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-213-888-6

Query Match 54.9%; Score 84; DB 9; Length 545;  
Best Local Similarity 46.4%; Pred. No. 5.6e-05;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 2 GHTGSVLCLOYDERVITITGSSDSTVRYW 29  
Db 292 GHTSTVRCMHLEKRVVSGSRDATTLRW 324

RESULT 10  
US-09-328-877A-6

; Sequence 6, Application US/09328877A  
; Patent No. US20020177187A1  
; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Li, Jinhe

; APPLICANT: Pauley, Adele M.

; APPLICANT: Pharmacia & Upjohn Company

; TITLE OF INVENTION: Encode Them

; FILE REFERENCE: 6142

; CURRENT APPLICATION NUMBER: US/09/328, 877A

; CURRENT FILING DATE: 1999-06-09

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 6

; LENGTH: 545

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-328-877A-6

Query Match 54.9%; Score 84; DB 9; Length 545;  
Best Local Similarity 46.4%; Pred. No. 5.6e-05;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 2 GHTGSVLCLOYDERVITITGSSDSTVRYW 29  
Db 292 GHTSTVRCMHLEKRVVSGSRDATTLRW 324

RESULT 11  
US-09-213-888-5

; Sequence 5, Application US/09213888A  
; Patent No. US20020164683A1  
; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-5

Query Match  
Best Local Similarity 54.9%; Score 84; DB 9; Length 553;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHTGSVLCLOYDERVITGSSDSTVRW 29  
DB 305 GHTSTVRCMHLHEKRVVSGSRDRTLRLW 332

RESULT 12  
US-09-328-877A-5  
Sequence 5, Application US/09328877A  
Patent No. US2002017187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Pharmacla & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-5

Query Match  
Best Local Similarity 54.9%; Score 84; DB 9; Length 553;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHTGSVLCLOYDERVITGSSDSTVRW 29  
DB 305 GHTSTVRCMHLHEKRVVSGSRDRTLRLW 332

RESULT 13  
US-09-213-888-9  
Sequence 9, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacla & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9

LENGTH: 559  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-9

Query Match  
Best Local Similarity 54.9%; Score 84; DB 9; Length 559;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHTGSVLCLOYDERVITGSSDSTVRW 29  
DB 311 GHTSTVRCMHLHEKRVVSGSRDRTLRLW 338

RESULT 14  
US-09-328-877A-9  
Sequence 9, Application US/09328877A  
Patent No. US2002017187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Pharmacla & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 559  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-9

Query Match  
Best Local Similarity 54.9%; Score 84; DB 9; Length 559;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHTGSVLCLOYDERVITGSSDSTVRW 29  
DB 311 GHTSTVRCMHLHEKRVVSGSRDRTLRLW 338

RESULT 15  
US-09-213-888-8  
Sequence 8, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacla & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 589  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-8

Query Match  
Best Local Similarity 54.9%; Score 84; DB 9; Length 589;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHTGSVLCLOYDERVITGSSDSTVRW 29  
DB 311 GHTSTVRCMHLHEKRVVSGSRDRTLRLW 338

Fri Apr 11 13:31:09 2003

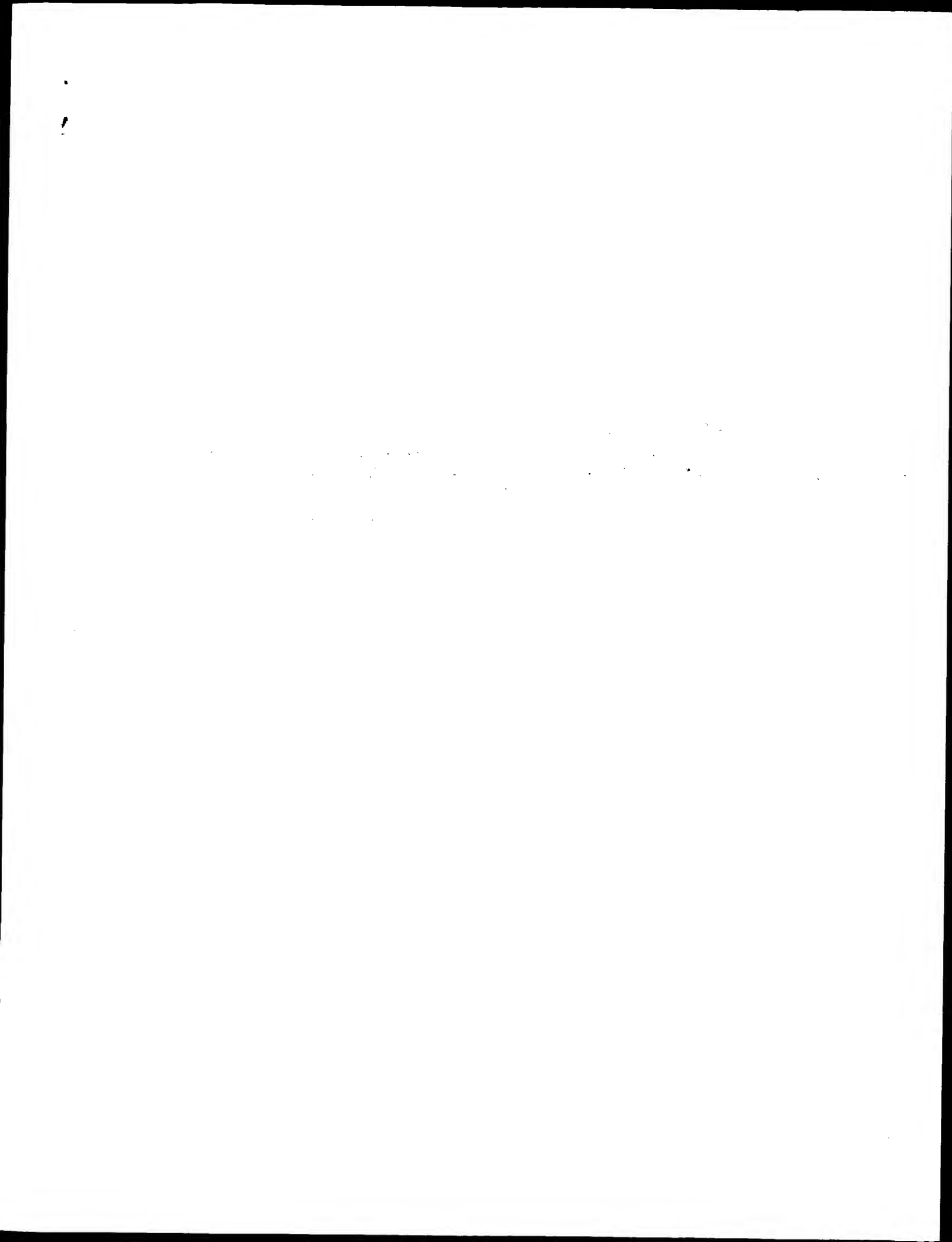
us-09-601-168b-2\_copy\_304\_332.rapb

Page 5

Db 341 GMTSTVRCMHLHEKRVVSGSRDATLRW 368

Search completed: April 10, 2003, 13:16:44  
Job time : 3.96711 secs

---



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 3.35566 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_304\_332  
Perfect score: 153  
Sequence: 1 TGTGTVLCIQYDERVITGSSDSTVRVW 29

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pIR1:  
2: pIR2:  
3: pIR3:  
4: pIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	148	96.7	518	2 B48088	beta-transducin re
2	137	89.5	701	2 T16607	hypothetical prote
3	96	62.7	605	2 T38932	probable sulfur me
4	90	58.8	267	2 S62507	hypothetical trp-a
5	90	58.8	506	2 T50211	WD-repeat protein
6	89	58.2	650	2 T46660	sulfur controller-
7	86	56.2	640	2 S49932	MEF30 protein - ye
8	83.5	54.6	1446	2 T13018	hypothetical prote
9	83	54.2	714	2 T22703	hypothetical prote
10	81	52.9	579	2 T22703	hypothetical prote
11	81	52.9	659	2 S38108	hypothetical prote
12	78	51.0	325	2 S60335	hypothetical prote
13	77.5	50.7	473	2 T33805	hypothetical prote
14	76	49.7	715	2 S38051	hypothetical prote
15	76	49.7	1189	2 A12493	DOA1 protein - ye
16	76	49.7	1258	2 A12155	WD-repeat protein
17	75.5	49.3	519	2 A49367	transducin homolog
18	75	49.0	515	2 S19487	hypothetical prote
19	75	49.0	558	2 AE2415	WD-repeat protein
20	75	49.0	651	2 T50289	WD-repeat protein
21	75	49.0	926	2 G95653	probable coatomer
22	74	48.4	333	2 G85034	probable WD-repeat
23	74	48.4	438	2 T45823	hypothetical prote
24	74	48.4	504	2 T50983	probable pleiotrop
25	74	48.4	559	2 AB2202	hypothetical prote
26	73.5	48.0	404	2 T40553	trp-asp repeat pro
27	73.5	48.0	779	2 S56245	cell division cont
28	73	47.7	422	2 A56640	CDC4 repeat unit-c
29	73	47.7	906	2 S35342	Golg1-associated p

30	73	47.7	906	2 S35312	coatomer complex b
31	73	47.7	1224	1 ERH04H	coatomer complex a
32	72.5	47.4	593	2 E96526	hypothetical prote
33	72	47.1	316	2 S57839	CPC2 protein - Neu
34	72	47.1	356	2 T22478	hypothetical prote
35	72	47.1	437	2 S05357	hypothetical prote
36	72	47.1	502	2 T41148	trp-asp repeat con
37	72	47.1	704	2 S33263	transcription init
38	72	47.1	755	2 T00066	hypothetical prote
39	72	47.1	1019	2 JC7538	neuronal different
40	72	47.1	1151	2 A55532	myosin-heavy-chain
41	72	47.1	1151	2 T33777	hypothetical prote
42	71.5	46.7	495	2 T47172	hypothetical prote
43	71	46.4	317	2 T46032	WD-40 repeat regul
44	71	46.4	328	2 A84901	hypothetical prote
45	71	46.4	589	2 AG2400	WD-repeat protein

## ALIGNMENTS

## RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-trcp

C:Species: Xenopus laevis (African clawed frog)

C>Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000

C:Accession: B48088

R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: A48088; MUID:93330289; PMID:8393141

A:Accession: B48088

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SPE>

A:Cross-references: GB:M98268; NID:q295542; PIDN:AAA02810.1; PID:q295543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <WD1>

Query Match 96.7%; Score 148; DB 2; Length 518;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHTGTVLCIQYDERVITGSSDSTVRVW 29  
DB 270 GHTGTVLCIQYDERVITGSSDSTVRVW 297

RESULT 2  
T16607 hypothetical protein K10B2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16607

R:Miller, N.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid K10B2.

A:Reference number: Z18545

A:Accession: T16607

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-701 <ML>

A:Cross-references: EMBL:U28730; NID:q860694; PID:q860695; PIDN:AAA68258.1; CESP:K10B

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K10B2.1

A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 89.5%; Score 137; DB 2; Length 701;  
Best Local Similarity 86.2%; Pred. No. 9.8e-13;  
Matches 25; Conservative 3; Mismatches 1; Indels 0; Gaps 0;





Db 543 GHTDGIITSLKFDSEKLVTVGSMDSVRIM 570

# RESULT 12

S60335

TGF-beta receptor interacting protein 1 - human

C:Species: Homo sapiens (man)

C>Date: 14-Sep-1996 #sequence\_revision 27-Feb-1997 #text\_change 28-May-1999

C:Accession: S60335

R:Chen, R.H.; Miettinen, P.J.; Maruoka, E.M.; Choy, L.; Derynck, R.

Nature 377, 548-552, 1995

A:Title: A WD-domain protein that is associated with and phosphorylated by the type II T

A:Reference number: S60256, PMID:96013749, PMID:7566156

A:Accession: S60335

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Cross-references: GB:U36764; NID:q1036804; PIDN:AAC50224.1; PID:q1036805

C:Superfamily: TGF-beta receptor interacting protein, WD repeat homology

F:184-217/Domain: WD repeat homology <MD1>

## Query Match

Best Local Similarity 51.0%; Score 78; DB 2; Length 325;

Matches 13; Conservative 10; Mismatches 5; Indels 2; Gaps 1;

Db 50 GHTGAVWCVDADMDTRHVLTVGSDNCRIM 79

# RESULT 13

hypothetical protein W07E6.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 26-May-2000

C:Accession: T33805

R:Latreille, P.; Mamsley, P.

submitted to the EMBL Data Library, November 1998

A:Description: The sequence of C. elegans cosmid W07E6.

A:Reference number: 221414

A:Accession: T33805

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-473 <LAT>

A:Cross-references: EMBL:AF106576; PIDN:AAC78176.1; GSPDB:GN00020; CESP:W07E6.2

A:Experimental source: strain Bristol N2; clone W07E6

C:Genetics:

A:Gene: CESP:W07E6.2

A:Map position: 2

A:Introns: 46/2; 77/3; 103/1; 195/2; 256/3; 311/3; 399/2

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 50.7%; Score 77.5; DB 2; Length 473;

Best Local Similarity 50.0%; Pred. No. 0.00085;

Matches 15; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

Db 234 SGHTASVTLRWGEGLYSGSDRTVKWM 263

# RESULT 14

S38051

DNAI protein - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 26-May-2000

C:Accession: S38051; S38056; S44325; S46544; S61942; S71950

R:Pohl, T.M.; Pohl, F.M.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S37897

A:Accession: S38051

A:Molecule type: DNA

A:Residues: 1-715 <POH>

A:Cross-references: EMBL:Z28213; NID:9486380; PIDN:CAA82058.1; PID:9486381; MIPS:YKL2

A:Experimental source: strain S288C

R:Alexandraki, D.; Tzermita, M.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S38054

A:Accession: S38056

A:Molecule type: DNA

A:Residues: 1-715 <ALE>

A:Cross-references: EMBL:Z28213; NID:9486380; PIDN:CAA82058.1; PID:9486381; MIPS:YKL2

A:Experimental source: strain S288C

R:Tzermita, M.; Horaitis, O.; Alexandraki, D.

Yeast 10, 663-679, 1994

A:Title: The complete sequencing of a 24.6 kb segment of yeast chromosome XI identifi

e dehydratases, membrane transporters, hydrolases and the phospholipase A(2)-activ

A:Reference number: S44319; MUID:95028164; PMID:7941750

A:Accession: S44325

A:Molecule type: DNA

A:Residues: 1-345 <TZE>

A:Cross-references: EMBL:X75951

R:Alexandraki, D.

submitted to the EMBL Data Library, December 1993

A:Reference number: S43546

A:Accession: S46544

A:Molecule type: DNA

A:Residues: 1-715 <ALM>

A:Cross-references: EMBL:X75951; NID:9473130; PIDN:CAA53560.1; PID:9473137

R:Hochstrasser, M.; Gang, G.

submitted to the EMBL Data Library, November 1995

A:Reference number: S61941

A:Accession: S61942

A:Molecule type: DNA

A:Residues: 1-715 <HOC>

A:Cross-references: EMBL:U39947; NID:q1086568; PIDN:AAA82258.1; PID:q1086570

R:Ghislain, M.; Dohmen, R.J.; Levy, F.; Varshavsky, A.

EMBO J. 15, 4884-4899, 1996

A:Title: Cdc48p interacts with Ufd3p, a WD repeat protein required for ubiquitin-medi

A:Reference number: S71950; MUID:97045097; PMID:8690162

A:Accession: S71950

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-715 <GHI>

C:Genetics:

A:Gene: SGD:DOA1; UFD3

A:Cross-references: SGD:S0001696; MIPS:YKL213c

A:Map position: 11L

C:Function:

A:Description: required for ubiquitin-dependent proteolysis

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

F:95-126/Domain: WD repeat homology <MD1>

F:133-167/Domain: WD repeat homology <MD2>

Query Match 49.7%; Score 76; DB 2; Length 715;

Best Local Similarity 46.4%; Pred. No. 0.0023;

Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Db 97 GHGNCVSLSFQDGVIVISGSMKTKAKVM 124

# RESULT 15

A12493

WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: A12493

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

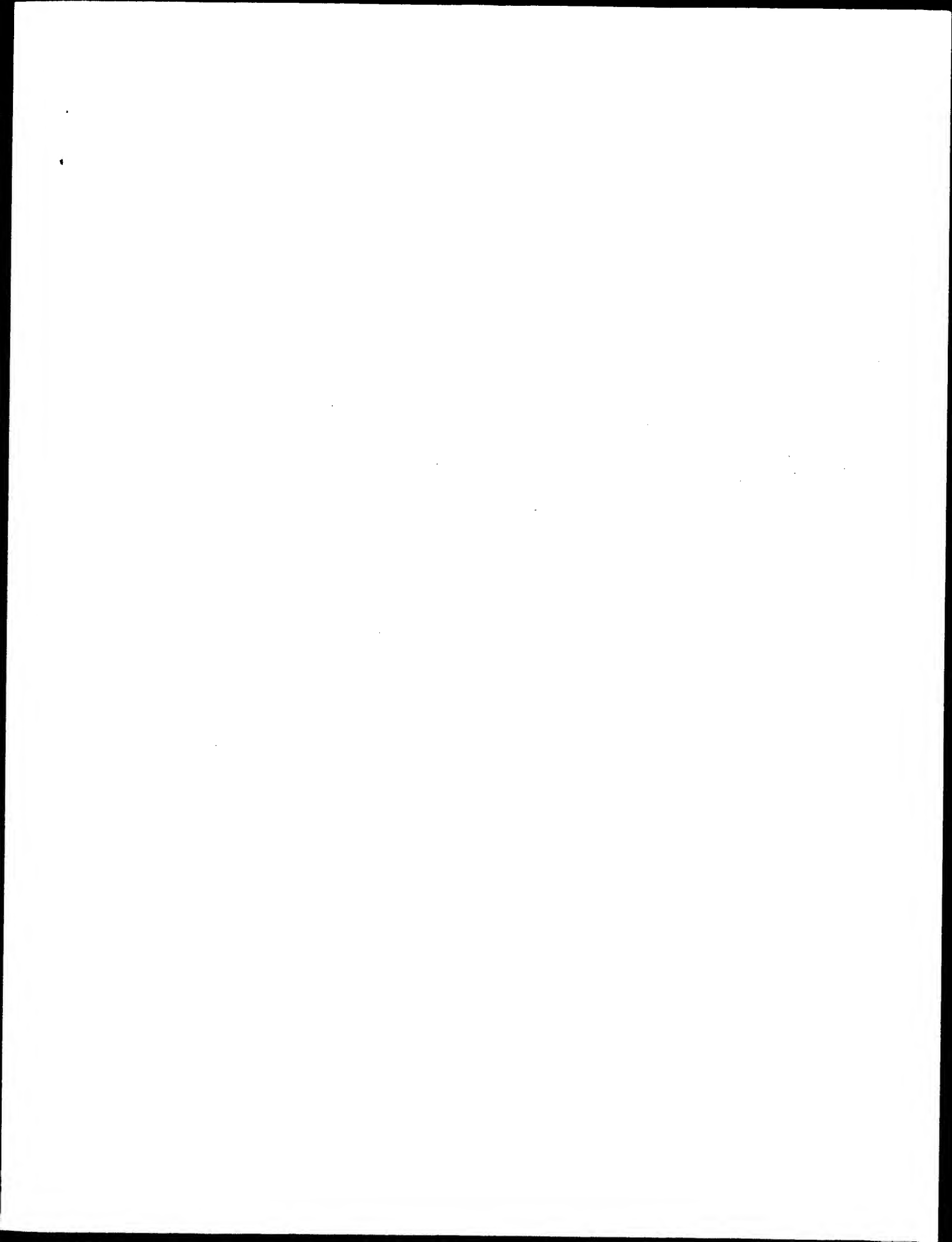
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium



A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: A12493  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1189 <KUR>  
A:Cross-references: GB:BA000020; PIDN:BAB78213.1; PID:917135667; GSEDB:GN00180  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr7129  
A:Genome: plasmid

Query Match	49.7%	Score 76	DB 2	Length 1189
Best Local Similarity	41.9%	Pred. No. 0.0042		
Matches 13	Conservative	9	Mismatches 7	Indels 2
				Gaps 1
QY	1	TGHTGCVLCLOY--DERVIITGSSDSTRYRW	29	
	:		:	:
	:		:	:
	:		:	:
	:		:	:
Db	1072	SGHTSOVCILFTKDKGRMRMISSSRITKIM	1102	

Search completed: April 10, 2003, 13:23:01  
Job time : 4.35566 secs



GenCore version 5.1.4-P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:51:26 ; Search time 2.08404 Seconds  
(without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_304\_332

Perfect score: 153  
Sequence: 1 TGHGSLVLCQYDERVITGSSDSTVRVW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	100.0	605	1	FWIA_HUMAN
2	152	99.3	542	1	FWIB_HUMAN
3	148	96.7	518	1	TRCB_XENLA
4	137	89.5	665	1	LIT3_CAEEL
5	98	64.1	678	1	SCOB_EMRNT
6	96	62.7	605	1	POFI_SCHPO
7	90	58.8	506	1	POFB_SCHPO
8	89	58.2	650	1	SCO2_NEUCR
9	86	56.2	640	1	MT30_YEAST
10	85	55.6	684	1	CC4_CANAL
11	83	54.2	714	1	YJ12_YEAST
12	81	52.9	579	1	SE10_CAEEL
13	79	51.6	659	1	YK16_YEAST
14	79	51.6	422	1	FBW2_HUMAN
15	78	51.0	325	1	IF32_HUMAN
16	78	51.0	325	1	IF32_MOUSE
17	76	49.7	326	1	IP32_DROME
18	76	49.7	715	1	DOAL_YEAST
19	76	49.7	1258	1	YS00_ANASP
20	75.5	49.3	519	1	TBL3_HUMAN
21	75	49.0	515	1	YCW2_YEAST
22	75	49.0	651	1	YX11_SCHPO
23	74	48.4	1224	1	COPA_BOVIN
24	73.5	48.0	779	1	GC4_YEAST
25	73	47.7	332	1	GRIP_DICDI
26	73	47.7	422	1	FBW2_MOUSE
27	73	47.7	905	1	COPB_BOVIN
28	73	47.7	905	1	COPB_HUMAN
29	73	47.7	905	1	COPB_MOUSE
30	73	47.7	1224	1	COPA_HUMAN
31	72	47.1	316	1	GRIP_NEUCR
32	72	47.1	356	1	GBB2_CAEEL
33	72	47.1	437	1	AAC3_DICDI

34	72	47.1	704	1	T2D4_DROME
35	72	47.1	904	1	COPB_RAT
36	72	47.1	1146	1	KHMA_DICDI
37	71.5	46.7	474	1	CO1C_HUMAN
38	71.5	46.7	474	1	CO1C_MOUSE
39	71	46.4	317	1	GRIP_HYDAT
40	71	46.4	775	1	POPI_SCHPO
41	71	46.4	1261	1	APAF_BRARE
42	70.5	46.1	425	1	RBH4_HUMAN
43	70.5	46.1	425	1	RBH7_HUMAN
44	70.5	46.1	425	1	RBH7_MOUSE
45	70.5	46.1	430	1	CAPI_DROME

## ALIGNMENTS

RESULT 1	FWIA_HUMAN	STANDARD	PRT	605 AA.
ID	FWIA_HUMAN			
AC	09Y297; 09Y213;			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TrCP)			
DE	(E3RSIKappab) (pikappaalapha-E3 receptor subunit).			
GN	FBXW1A OR FBW1A OR BTCP OR BTCP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=99075339; PubMed=9859996;			
RA	Varon A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,			
RA	Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;			
RT	"Identification of the receptor component of the IkappaBalpha-			
RT	ubiquitin ligase."			
RL	Nature 396:590-594(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=lymphoid;			
RX	MEDLINE=96325370; PubMed=9660940;			
RA	Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,			
RA	Thomas D., Strebel K., Benarous R.;			
RT	"A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu			
RT	connects CD4 to the ER degradation pathway through an F-box motif."			
RL	Mol. Cell 1:565-574(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RX	MEDLINE=20003060; PubMed=10531035;			
RA	Cenciarelli C., Chlaud D.S., Guardavaccaro D., Parks W., Vidal M.,			
RA	Pagan M.;			
RT	"Identification of a family of human F-box proteins."			
RL	Curr. Biol. 9:1177-1179(1999).			
RN	[4]			
RP	CHARACTERIZATION.			
RX	MEDLINE=99145464; PubMed=9990852;			
RA	Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,			
RA	Harper J.W.;			
RT	"The SCF(beta-TrCP)-ubiquitin ligase complex associates specifically			
RT	with phosphorylated destruction motifs in I-kappa-B-alpha and			
RT	beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro."			
RL	Genes Dev. 13:270-283(1999).			
CC	- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA			
CC	(PIKAPPALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR			
CC	UBQUITINATION AND DEGRADATION.			
CC	- SUBUNIT: PART OF A SCF (SKP1-CULIN-F-BOX) PROTEIN LIGASE COMPLEX.			
CC	- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE			
CC	PRODUCED BY ALTERNATIVE SPLICING.			
CC	- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.			
CC	- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF101784; AAD08702.1;  
 DR EMBL: Y14153; CAZ74572.1;  
 DR EMBL: AF129530; AAF04464.1;  
 DR Genew: HGNC:1144; BTRC.  
 DR MIM: 603482;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 KW DOMAIN 190 228 F-BOX.  
 FT REPEAT 301 338 WD 1.  
 FT REPEAT 341 378 WD 2.  
 FT REPEAT 381 418 WD 3.  
 FT REPEAT 424 461 WD 4.  
 FT REPEAT 464 503 WD 5.  
 FT REPEAT 505 541 WD 6.  
 FT REPEAT 553 590 WD 7.  
 FT VARSPLIC 17 52 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 605 AA; 68866 MW; 4C673B7E400F037 CRC64;

Query Match 100.0%; Score 153; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 1; 7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGHGTVLCLOYDERVITIGSSDSSTVRW 29  
 DB 340 TGHGTVLCLOYDERVITIGSSDSSTVRW 368  
 |||||||||||||||||||||||||||||

RESULT 2

FWIB\_HUMAN STANDARD; PRT; 542 AA.  
 AC Q90K81; Q9Y4C6; Q9P2S8; Q9P2S9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein 1b (F-box and WD-repeats protein beta-Trop2).  
 GN FBXW1B OR FBW1B OR BTBKP2 OR KIAA0696.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NC NCB1; TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20003060; PubMed=10531035;  
 RA Genclarelili C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins."  
 RL Curr. Biol. 9:1177-1179(1999).  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RC TISSUE=fetal lung;  
 RA MEDLINE=20160458; PubMed=10694485;  
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,  
 RA Katoh M.;

RT "Molecular Cloning and genomic structure of the betaTRCP2 gene on  
 RT chromosome 5q35.1.";  
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A, B AND C (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF176022; AAF04528.1; -;  
 DR EMBL: AB033279; BAA92329.1; -;  
 DR EMBL: AB033280; BAA92330.1; -;  
 DR EMBL: AB033281; BAA92331.1; -;  
 DR EMBL: AB014596; BAA31671.1; ALT\_INTT.  
 DR Genew: HGNC:13607; FBXW1B.  
 DR MIM: 605651; -;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 KW DOMAIN 129 167 F-BOX.  
 FT REPEAT 238 275 WD 1.  
 FT REPEAT 278 315 WD 2.  
 FT REPEAT 318 355 WD 3.  
 FT REPEAT 361 398 WD 4.  
 FT REPEAT 401 440 WD 5.  
 FT REPEAT 442 478 WD 6.  
 FT REPEAT 490 527 WD 7.  
 FT VARSPLIC 16 49 MISSING (IN ISOFORM A).  
 FT VARSPLIC 16 48 MISSING (IN ISOFORM B).  
 SQ SEQUENCE 542 AA; 62090 MW; 7CD40087EFAA5C8A CRC64;

Query Match 99.3%; Score 152; DB 1; Length 542;  
 Best Local Similarity 96.6%; Pred. No. 2; 2e-15;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGHGTVLCLOYDERVITIGSSDSSTVRW 29  
 DB 277 TGHGTVLCLOYDERVITIGSSDSSTVRW 305  
 |||||||||||||||||||||||||||||

RESULT 3

TRCB\_XENIA STANDARD: PRT: 518 AA.  
 ID TRCB\_XENIA PRT: 518 AA.  
 AC 091854; P70037; P70038;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Beta-TRCP (beta-transducin repeat-containing protein).  
 GN FAXM1 OR BTRCP.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NX NCBI\_TaxId=6355;  
 RX MEDLINE=9330289; PubMed=8393141;  
 RA Shevak W., Keiper B.D., Stralowa C., Castanon M.J.;  
 RT "Saccharomyces cerevisiae dccl5 mutants arrested at a late stage in  
 anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein  
 with beta-transducin repeats";  
 RL Mol. Cell. Biol. 13:4953-4966(1993).  
 RN [2]  
 RP SEQUENCE OF 302-518 FROM N.A.  
 RX MEDLINE=97109804; PubMed=8952061;  
 RA Hudson J.W., Alarcon V.B., Elinson R.P.;  
 RT "Identification of new localized RNAs in the Xenopus oocyte by  
 differential display PCR";  
 RL Dev. Genet. 19:190-198(1996).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX  
 (BY SIMILARITY).  
 CC -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-  
 MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC  
 GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO  
 NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR  
 TADPOLE EMBRYO.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG. TO C.ELEGANS K10B2.1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M98268; AAA02810.1; -;  
 DR EMBL: U63921; AAB49671.1; -;  
 DR EMBL: U63922; AAB49672.1; -;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001810; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00446; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ub1 conjugation pathway. Repeat: WD repeat.  
 FT DOMAIN 119 157 F-BOX.  
 FT REPEAT 230 258 WD 1.  
 FT REPEAT 270 298 WD 2.  
 FT REPEAT 310 338 WD 3.  
 FT REPEAT 353 381 WD 4.  
 FT REPEAT 393 421 WD 5.  
 FT REPEAT 433 461 WD 6.  
 FT REPEAT 482 510 WD 7.

FT CONFLICT 302 304 GEM -> EFR (IN REF. 2).  
 FT CONFLICT 516 518 GLA -> AAH (IN REF. 2).  
 FT SEQUENCE 518 AA: 59507 MW: 2A52EC19028127F3 CRC64:  
 Query Match 96.7%; Score 148; DB 1; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-15;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GHTGSVLCQYDERVITIGSSDSVTRVW 29  
 DB 270 GHTGSVLCQYDERVITIGSSDSVTRVW 297  
 RESULT 4  
 LI23\_CAEL STANDARD: PRT: 665 AA.  
 ID LI23\_CAEL PRT: 665 AA.  
 AC 009990; 096GN6;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein lin-23.  
 GN LIN-23 OR K10B2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodidae; Caenorhabditis.  
 NX NCBI\_TaxId=6239;  
 RX MEDLINE=20515608; PubMed=11060233;  
 RA Kipreos E.T., Gohei S.P., Hedgecock E.M.;  
 RT "The Caenorhabditis elegans F-box/WD-repeat protein lin-23 functions  
 to limit cell division during development";  
 RL Development 127:5071-5082(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RA Miller N.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Functions cell autonomously to negatively regulate cell  
 cycle progression. Required to restrain cell proliferation in  
 response to developmental cues. Probably recognizes and binds to  
 some proteins and promotes their ubiquitination and degradation  
 (By similarity).  
 CC -1- SUBUNIT: Part of a SCF (SKP1-CULLIN-F-BOX) protein ligase complex  
 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- DEVELOPMENTAL STAGE: Highest levels in embryos and adults, lowest  
 levels in larvae. Maternal expression results in high zygotic  
 levels.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG. TO X.LAEVIS FBXW1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF275253; AAG28037.1; -;  
 DR EMBL: U28730; AAG68258.2; -;  
 DR WormPeD: K10B2.1; CE28600.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001810; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00446; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.



```

RL Nature 415:871-880(2002).
CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY)
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB032410; BAA84528.1; -
DR EMBL: Z94864; CAB08168.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF004400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 5.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Ubl conjugation; Repeat; WD repeat.
KW DOMAIN
FT REPEAT 107 153 F-BOX.
FT REPEAT 271 299 WD 1.
FT REPEAT 311 339 WD 2.
FT REPEAT 350 379 WD 3.
FT REPEAT 390 420 WD 4.
FT REPEAT 432 460 WD 5.
FT REPEAT 472 500 WD 6.
FT REPEAT 510 538 WD 7.
SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 62.7%; Score 96; DB 1; Length 605;
Best Local Similarity 53.6%; Pred. No. 8e-07;
Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 GHTGSVLCLOYDERVITIGSSDSSTVRW 29
DB 351 GHTDSVLCITFDSTLVSGSADCTVKLM 378

RESULT 7
POPB_SCHPO STANDARD; PRT; 506 AA.
AC 009855; 0987V1;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein profil.
GN POP11 OR SPAC29E6.01 OR SPAC30.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrison C.L., Toda T.;
RT "Systematic genome-wide analysis of F-box protein-encoding genes in
RT fission yeast."
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=9712;

```

```

RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gebel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Mambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Diano S., Gloux S., Lelaue V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipkovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD_REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB061694; BAB55543.1; -
DR EMBL: AL136538; CAB66464.1; -
DR EMBL: Z66525; CAA91423.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF004400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 3.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 8.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT REPEAT 70 116 F-BOX.
FT REPEAT 219 256 WD 1.
FT REPEAT 259 298 WD 2.
FT REPEAT 301 338 WD 3.
FT REPEAT 345 386 WD 4.
FT REPEAT 388 426 WD 5.
FT REPEAT 427 464 WD 6.
FT REPEAT 468 505 WD 7.
SQ SEQUENCE 506 AA; 58257 MW; CEF34DAEFFBC2E10 CRC64;

Query Match 58.8%; Score 90; DB 1; Length 506;
Best Local Similarity 53.3%; Pred. No. 5.3e-06;
Matches 16; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

QY 2 GHTGSVLCLOYDER-VITIGSSDSSTVRW 29
DB 259 GHTGSVLCIDFCRRNLIVSGSSDSSTIIIM 288

```

```

RESULT 8
SC02_NEUCR          STANDARD:          PRT:          650 AA.
ID SC02_NEUCR
AC 001277;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sulfur controller-2 (SCON2).
GN SCON-2.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariiales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RX MEDLINE=95241499; PubMed=7724564;
RA Kumar A., Paletta J.V.;
RT "The sulfur controller-2 negative regulatory gene of Neurospora
RL crassa encodes a protein with beta-transducin repeats."
RC Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
CC -!- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC -!- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 017251; AAA68968.1; -
DR InterPro: IPR001810; F-box.
DR Pfam: PF00400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00320; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Repeat; WD repeat.
KM DOMAIN 124 170
FT REPEAT 292 320 WD 1.
FT REPEAT 332 360 WD 2.
FT REPEAT 372 400 WD 3.
FT REPEAT 411 441 WD 4.
FT REPEAT 453 488 WD 5.
FT REPEAT 528 564 WD 6.
FT REPEAT 576 604 WD 7.
FT REPEAT 616 644 WD 8.
SO SEQUENCE 650 AA; 72189 MW; 7473859C9P9F1B028 CRC64;

Query Match 58.28; Score 89; DB 1; Length 650;
Best Local Similarity 46.48; Pred. No. 1e-05;
Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

```

```

AC p39014;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MET30 protein.
GN MET30 OR YIL046W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=96069360; PubMed=8524217;
RA Thomas D., Kuras L., Barbey R., Charest H., Blaiseau P.L.,
RA Surlin-Kerjan Y.;
RT "Met30p, a yeast transcriptional inhibitor that responds to S-
RL Mol. Cell. Biol. 15:6526-6534(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Bartell B.G., Badcock R., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagsels K., Jones M.,
RA Louis E., Lye G., Moule S., Odell C., Pearson D.,
RA Raeburn M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (Dec-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
CC GENES EXPRESSION.
CC -!- SUBUNIT: SEEMS TO INTERACT WITH MET4.
CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z46861; CA86905.1; -
DR EMBL: L26505; AAA96717.1; -
DR SCD: S0001308; MET30.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 6.
DR PROSITE: PS00181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Methionine biosynthesis;
KW Cysteine biosynthesis; Repeat; WD repeat.
KM DOMAIN 181 227
FT REPEAT 300 328 WD 1.
FT REPEAT 340 368 WD 2.
FT REPEAT 380 408 WD 3.
FT REPEAT 419 449 WD 4.
FT REPEAT 461 499 WD 5.
FT REPEAT 509 538 WD 6.
FT REPEAT 550 578 WD 7.
FT REPEAT 607 635 WD 8 (POTENTIAL).
FT CONFLICT 61 61 M -> I (IN REF. 1).
SO SEQUENCE 640 AA; 72835 MW; 5135DABCA2E1EB97 CRC64;

```



Query Match 56.2%; Score 86; DB 1; Length 640;  
 Best Local Similarity 51.7%; Pred. No. 2.8e-05;  
 Matches 15; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 TGHTGSVLCLOYDERVITGSSDSTVRW 29  
 DB 339 SGHSDGVKLTLYFDRLKLTIGSLDKTRIRW 367

RESULT 10  
 CC4\_CANAL STANDARD: PRT; 684 AA.

AC P53699;  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cell division control protein 4.  
 GN CDC4.

OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=SGY126;  
 RA Shieh J.C., White A.M., Rosamond J.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA  
 REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE  
 POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD  
 DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND  
 VARIOUS ASPECTS OF SPOULATON. REQUIRED FOR HVAL-HTB1 LOCUS  
 TRANSCRIPTION ACTIVATION (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X96763; CA65538.1;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 1.  
 DR SMART: SM00320; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS00682; WD\_REPEATS\_2; 4.  
 DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
 KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.  
 DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
 FT DOMAIN 212 258 F-BOX.  
 FT REPEAT 322 351 WD 1.  
 FT REPEAT 363 391 WD 2.  
 FT REPEAT 403 431 WD 3.  
 FT REPEAT 442 468 WD 4 (POTENTIAL).  
 FT REPEAT 478 506 WD 5.  
 FT REPEAT 519 549 WD 6.  
 FT REPEAT 561 589 WD 7.  
 SO SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;

Query Match 55.6%; Score 85; DB 1; Length 684;  
 Best Local Similarity 53.6%; Pred. No. 4.4e-05;  
 Matches 15; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 GHTGSVLCLOYDERVITGSSDSTVRW 29  
 DB 363 GHEGVWALKTYGNTLVGSDTRVIRW 390

RESULT 11  
 YJL2\_YEAST STANDARD: PRT; 714 AA.

AC P47025;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 80.0 kDa Trp-Asp repeats containing protein in ASF1-CTT7  
 intergenic region.  
 GN YJL112W OR J0802.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / FY1679;  
 RX MEDLINE=97103775; PubMed=8948101;  
 RA Cziepluch C., Kordes E., Pujoil A., Jauniaux J.-C.;

RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X  
 reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,  
 SP10, CDC14, RPE1, PHO86, NCA3, ASF1, COT7, GZF3, two tRNA genes,  
 three remnant delta elements and a Ty4 transposon."  
 RL Yeast 12:1471-1474(1996).

RL [2]  
 RP SEQUENCE OF 596-714 FROM N.A.

RC STRAIN=S288C / FY1679;  
 RX MEDLINE=96090136; PubMed=7483851;  
 RA Rasmussen S.W.;

RT "A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1  
 and GCD3 genes, a TCP-1-related gene, an open reading frame similar  
 to the DAL80 gene, and a tRNA(Arg).";

RL Yeast 11:873-883(1995).  
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z49387; CA89407.1;  
 DR SCD: S0003648; YJL112W.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 2.  
 DR SMART: SM00320; WD40; 5.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS00682; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 396 427 WD 1.  
 FT REPEAT 439 469 WD 2.  
 FT REPEAT 500 528 WD 3.  
 FT REPEAT 564 592 WD 4.  
 FT REPEAT 604 632 WD 5.  
 FT REPEAT 643 672 WD 6.  
 SO SEQUENCE 714 AA; 80032 MW; CA8A23DB06193A6 CRC64;

Query Match 54.2%; Score 83; DB 1; Length 714;  
 Best Local Similarity 46.4%; Pred. No. 9.3e-05;  
 Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

```

AC P36130; 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 74.7 kDa Tfp-Asp repeats containing protein in DAL80-GAP1
DE Intergenic region.
DE YKR036C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932.
RN [1]
RP SEQUENCE FROM N.A.
RF Uffersatrazu L.A., Jauniaux J.-C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: CONTAINS 6 WD REPEATS (TFR-ASP DOMAINS).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC of the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/annouce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL; Z28261; CAA82110.1; -
CC FTR; S38108; S38108.
CC TRANSFAC; T04326; -.
CC
CC SGD; S0001744; YKR036C.
CC Interpro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 6.
CC PRINTS; PR00320; GPROTINBRPT.
CC PRODOM; PD000018; WD40; 1.
CC SMART; SM00320; WD40; 5.
CC PROSITE; PS00678; WD_REPEATS_1; 4.
CC PROSITE; PS0082; WD_REPEATS_2; 5.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
FT REPEAT 333..364 WD 1.
FT REPEAT 376..407 WD 2.
FT REPEAT 438..466 WD 3.
FT REPEAT 503..531 WD 4.
FT REPEAT 543..571 WD 5.
FT REPEAT 582..610 WD 6.
SQ SEQUENCE 659 AA; 74709 MW; 139593554C56DD37 CRC64;

Query Match 52.9%; Score 81; DB 1; Length 659;
Best Local Similarity 42.9%; Pred. No. 0.00017;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

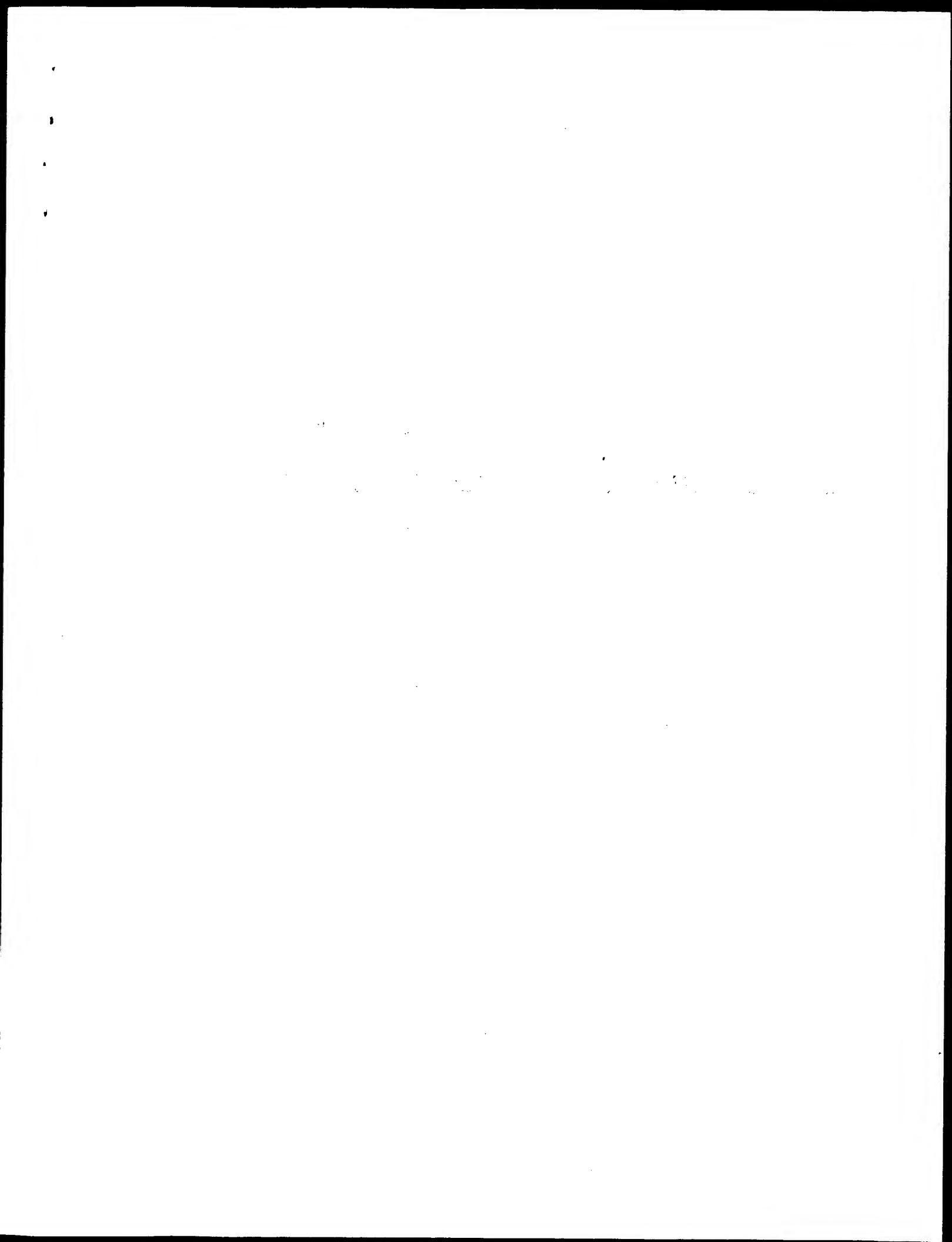
QY 2 GHTGSVLCTQYDERVILINGSSDSPTVRYW 29
DB 543 GHTDGTSLKFDSEKLVGSMDSVRYW 570

RESULT 14
FBW2_HUMAN
ID FBW2_HUMAN STANDARD; PRT; 422 AA.
AC O9UKR8; O9UKR3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein 2.
DE FBW2 OR FBW2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606.
RN [1]
RP SEQUENCE FROM N.A.
RF MEDLINE=20003060; PubMed=10531035;
RA Genclarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,
RA Pagano M.;

```

RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2003061; PubMed=10531037;  
 RA Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.;  
 RT "A family of mammalian F-box proteins.";  
 RL Curr. Biol. 9:1180-1182(1999).  
 CC -1- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHORYLATED  
 CC PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-COLLIN-F-BOX) PROTEIN LIGASE COMPLEX  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF129531; AAF04465.1; -;  
 DR EMBL; AF176698; AAF13226.1; -;  
 DR Genew; HGNC:13608; FBXW2.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00440; WD40; 4.  
 DR Pfam; PF00646; F-box; 1.  
 DR PRINTS; PR00320; GPOTEINBRPT.  
 DR PRODOM; PD000018; WD40; 1.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 4.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ubl conjugation pathway; Repeat; WD repeat.  
 KW DOMAIN 54 101 F-BOX.  
 FT REPEAT 146 183 WD 1.  
 FT REPEAT 185 221 WD 2.  
 FT REPEAT 224 265 WD 3.  
 FT REPEAT 276 314 WD 4.  
 FT CONFLICT 408 408 E -> K (IN REF. 2).  
 FT SEQUENCE 422 AA; 47966 MW; 2874BARF233DD8E CRC64;  
 SQ  
 Query Match 51.6%; Score 79; DB 1; Length 422;  
 Best Local Similarity 43.3%; Pred. No. 0.00021;  
 Matches 13; Conservative 10; Mismatches 5; Indels 2; Gaps 1;  
 Oy 2 GHTGVCVLCYDQDER--VITGSSDSTVRVW 29  
 Db 224 GHTGAVFSVDYNDLIDLVGSADPTVAVW 253  
 RESULT 15  
 IF32\_HUMAN STANDARD: PRT: 325 AA.  
 ID IF32\_HUMAN STANDARD: PRT: 325 AA.  
 AC Q13347;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Eukaryotic translation initiation factor 3 subunit 2 (eIF-3 beta)  
 DE (eIF3 p36) (TGF-beta receptor interacting protein 1) (TRIP-1).  
 GN EIF3S2 OR TRIP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;

RX MEDLINE=97150873; PubMed=8995409;  
 RA Asano K., Kinzy T.G., Merrick W.C., Hershey J.W.B.;  
 RT "Conservation and diversity of eukaryotic translation initiation  
 RT factor eIF3.";  
 RL J. Biol. Chem. 272:1101-1109(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96013749; PubMed=7566156;  
 RA Chen R.H., Mittleman P.J., Maruka E.M., Choy L., Derynck R.;  
 RT "A chd domain protein that is associated with and phosphorylated by  
 RT the type II TGF-beta receptor.";  
 RL Nature 377:548-552(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung, and Muscle;  
 RA Strassberg R.;  
 RL Submitted (Feb-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF  
 CC METHIONYL-TRNAI AND MRNA.  
 CC -1- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS.  
 CC -1- PTM: PHOSPHORYLATED BY TYPE II TGF-BETA RECEPTOR.  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U39067; AAC97144.1; -;  
 DR EMBL; U36764; AAC50224.1; -;  
 DR EMBL; BC000413; AAH00413.1; -;  
 DR EMBL; BC003140; AAH03140.1; -;  
 DR PHCI-2DPAGE; Q13347; -;  
 DR Genew; HGNC:3272; EIF3S2.  
 DR MIM; 603911; -;  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00440; WD40; 5.  
 DR PRINTS; PR00320; GPOTEINBRPT.  
 DR PRODOM; PD000018; WD40; 2.  
 DR SMART; SM00320; WD40; 5.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 4.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 2.  
 DR Initiation factor; Protein biosynthesis; Repeat; WD repeat;  
 KW Phosphorylation.  
 FT REPEAT 8 38 WD 1.  
 FT REPEAT 50 80 WD 2.  
 FT REPEAT 144 174 WD 3.  
 FT REPEAT 186 216 WD 4.  
 FT REPEAT 283 313 WD 5.  
 FT SEQUENCE 325 AA; 36502 MW; 02797BB72A752A96 CRC64;  
 SQ  
 Query Match 51.0%; Score 78; DB 1; Length 325;  
 Best Local Similarity 43.3%; Pred. No. 0.00021;  
 Matches 13; Conservative 10; Mismatches 5; Indels 2; Gaps 1;  
 Oy 2 GHTGVCVLC--QYDERVITGSSDSTVRVW 29  
 Db 50 GHTGAVWCVDADWDTKHVLTSADNSCRWL 79  
 Search completed: April 10, 2003, 13:19:34  
 Job time : 3.08404 secs



GenCore version 5.1.4-P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 ; Search time 6.85262 Seconds  
(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_304\_332

Sequence: 1 TGHGSGVICLDYDERVITITGSSDSTRVW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.podent:\*
- 12: sp.virus:\*
- 13: sp.vertbrate:\*
- 14: sp.unclassified:\*
- 15: sp.rvirus:\*
- 16: sp.bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	153	100.0	569	11	Q9R1G7	Q9R1G7 mus musculus
2	153	100.0	569	11	Q9R0I5	Q9R0I5 mus musculus
3	152	99.3	563	11	Q923H0	Q923H0 mus musculus
4	146	95.4	569	11	Q92159	Q92159 mus musculus
5	138	90.2	510	5	Q44382	Q44382 drosophila
6	138	90.2	510	5	Q9VDE3	Q9VDE3 drosophila
7	124	81.0	38	13	Q8UUN4	Q8UUN4 xenopus lae
8	124	81.0	430	5	Q9B5J4	Q9B5J4 heterodera
9	97	63.4	252	11	Q922C7	Q922C7 mus musculus
10	89	58.2	280	5	Q8S7S5	Q8S7S5 encephalito
11	87.5	57.2	414	5	Q8SW59	Q8SW59 homo sapien
12	84	54.9	553	4	Q9N0X6	Q9N0X6 homo sapien
13	84	54.9	561	4	Q96R12	Q96R12 homo sapien
14	84	54.9	589	4	Q96LE0	Q96LE0 homo sapien
15	84	54.9	627	4	Q96A16	Q96A16 homo sapien
16	84	54.9	629	11	Q8VHP4	Q8VHP4 mus musculus

17	84	54.9	629	11	Q8VBP4	Q8VBP4 mus musculus
18	84	54.9	707	4	Q969H0	Q969H0 homo sapien
19	83.5	54.6	1446	10	Q9SUT4	Q9SUT4 arabidopsis
20	83	54.2	388	10	Q9FC99	Q9FC99 arabidopsis
21	82	53.6	1228	5	Q9U4H2	Q9U4H2 drosophila
22	82	53.6	1240	5	Q9V815	Q9V815 drosophila
23	81	52.9	585	5	Q952T0	Q952T0 caenorhabd
24	81	52.9	587	5	Q44083	Q44083 caenorhabd
25	81	52.9	1326	5	Q9VZF4	Q9VZF4 drosophila
26	80	52.3	122	13	Q8UUN3	Q8UUN3 xenopus lae
27	79	51.6	454	4	Q8WV51	Q8WV51 homo sapien
28	79	51.6	454	4	Q9HA09	Q9HA09 homo sapien
29	79	51.6	562	10	Q9SC73	Q9SC73 oryza sativ
30	79	51.6	922	5	Q9VE98	Q9VE98 drosophila
31	79	51.6	1204	10	Q8S0C3	Q8S0C3 oryza sativ
32	78	51.0	1241	2	Q9XBD8	Q9XBD8 amycollops
33	77.5	50.7	447	10	Q9AV17	Q9AV17 oryza sativ
34	77.5	50.7	473	5	Q9TVV3	Q9TVV3 caenorhabd
35	77	50.3	914	4	Q9CE93	Q9CE93 homo sapien
36	77	50.3	1573	11	Q90XL2	Q90XL2 mus musculus
37	76.5	50.0	613	5	Q960M2	Q960M2 drosophila
38	76.5	50.0	621	5	Q9VAT2	Q9VAT2 drosophila
39	76	49.7	307	11	Q8R315	Q8R315 mus musculus
40	76	49.7	367	4	Q9H8N9	Q9H8N9 homo sapien
41	76	49.7	415	4	Q9P2B3	Q9P2B3 h kiara135
42	76	49.7	992	16	Q98GJ0	Q98GJ0 rhizobium
43	76	49.7	1189	16	Q8Y109	Q8Y109 anabaena sp
44	75.5	49.3	808	4	Q96A78	Q96A78 homo sapien
45	75	49.0	304	11	Q9DAT2	Q9DAT2 mus musculus

## ALIGNMENTS

## RESULT 1

Q9R1G7 PRELIMINARY: PRT: 569 AA.

AC 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Beta-transducin repeat-containing protein.  
CN BTRC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.  
CX NCBI\_TaxID=10090;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Winston J., Elledge S.J., Harper J.W.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRF-ASP DOMAINS).  
DR EMBL: AF110396; AD41025.1; -.  
DR MGD: MGI:1338871; Btrc.  
DR InterPro: IPR001810; F-box.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00646; F-box; 1.  
DR Pfam: PF00400; WD40; 7.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR PRODOM: PD000018; WD40; 4.  
DR SMART: SM00256; FBOX; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS0181; PROX; 1.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
DR PROSITE: PS00082; WD\_REPEATS\_2; 7.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.

SEQUENCE 569 AA: E6DDCAD28D551D9D CRC64:

Query Match 100.0%; Score 153; DB 11; Length 569;  
Best Local Similarity 100.0%; Pred. No. 3.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHGTVLCLQYDERVITITGSSDSTVRW 29  
 DB 304 TGHGTVLCLQYDERVITITGSSDSTVRW 332

RESULT 2  
 ID 090015 PRELIMINARY; PRT; 569 AA.  
 AC 090015;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DE 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)  
 DE (F-box-WD40 repeat protein 1).  
 GN BTRC OR FBXW1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99199275; PubMed-10097128;  
 RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,  
 RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,  
 RA Nakayama K.-I.;  
 RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a  
 RT ubiquitin ligase Skp1/Cul1/F-box protein FWD1";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99075339; PubMed-9859996;  
 RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
 RT "Identification of the receptor component of the IkappaBalpha-  
 RT ubiquitin ligase";  
 RL Nature 396:590-594(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV;  
 RX MEDLINE-21601157; PubMed=11735228;  
 RA Matuyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,  
 RA Nakayama K.-I.;  
 RT "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of  
 RT Caenorhabditis elegans SEL-10.";  
 RL Genomics 78:214-222(2001).  
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF081887; AAD17755.1;  
 DR EMBL: AF099932; AAD08701.1;  
 DR EMBL: BC003989; AAH03989.1;  
 DR EMBL: AF391190; AAL40929.1;  
 DR EMBL: AF391178; AAL40929.1; JOINED.  
 DR EMBL: AF391179; AAL40929.1; JOINED.  
 DR EMBL: AF391180; AAL40929.1; JOINED.  
 DR EMBL: AF391181; AAL40929.1; JOINED.  
 DR EMBL: AF391182; AAL40929.1; JOINED.  
 DR EMBL: AF391183; AAL40929.1; JOINED.  
 DR EMBL: AF391184; AAL40929.1; JOINED.  
 DR EMBL: AF391185; AAL40929.1; JOINED.  
 DR EMBL: AF391186; AAL40929.1; JOINED.  
 DR EMBL: AF391187; AAL40929.1; JOINED.  
 DR EMBL: AF391188; AAL40929.1; JOINED.  
 DR EMBL: AF391189; AAL40929.1; JOINED.  
 DR MGD; MGI:133871; Btrc.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PRO0320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.

DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ligase; Repeat; WD repeat.  
 SO SEQUENCE 569 AA; 65105 MW; BC7D6544815B296 CRC64;

Query Match 100.0%; Score 153; DB 11; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3; 7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHGTVLCLQYDERVITITGSSDSTVRW 29  
 DB 304 TGHGTVLCLQYDERVITITGSSDSTVRW 332

RESULT 3  
 ID 0923H0 PRELIMINARY; PRT; 563 AA.  
 AC 0923H0;  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DE 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE F-box-WD40 repeat-containing protein HOS.  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N;  
 RA Bhatia N., Heter J.R., Slaga J.J., Fuchs S.Y., Spiegelman V.S.;  
 RT "Mouse homolog of HOS (rhOS) is overexpressed in skin tumors and  
 RT implicated in constitutive activation of NF-kappaB";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: A1038079; AAK72095.1;  
 DR MGD; MGI:2144023; Fbxw1b.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRODOM: PD000018; WD40; 4.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS500678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SO SEQUENCE 563 AA; 64741 MW; 9AB562F3FF5E3496 CRC64;

Query Match 99.3%; Score 152; DB 11; Length 563;  
 Best Local Similarity 96.6%; Pred. No. 5; 2e-15;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGHGTVLCLQYDERVITITGSSDSTVRW 29  
 DB 298 TGHGTVLCLQYDERVITITGSSDSTVRW 326

RESULT 4  
 ID 092159 PRELIMINARY; PRT; 569 AA.  
 AC 092159;  
 DT 01-MAY-1999 (TRENBLREL. 10, Created)  
 DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)  
 DE 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Beta-transducin repeat containing protein.  
 GN BTRC.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99145465; PubMed=9990853;  
 RA Spencer E., Jiang J., Chen Z.J.;  
 RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein  
 SLMB/Delta-TRCP.";  
 RT Genes Dev. 13:284-294(1999).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF112979; AAD04181.1; .  
 DR MGD: MGI:1338871; Btrc.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001880; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat: WD repeat.  
 SQ SEQUENCE 569 AA; 65047 MW; BCTG7A44815BED96 CRC64;  
 Query Match 95.4%; Score 146; DB 11; Length 569;  
 Best Local Similarity 96.6%; Pred. No. 4,5e-14;  
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 TGHTGSLCQYDERVITITSSSDSTVRW 29  
 Db 304 TGHTGSLCQYDERVITITSSSDSTVRW 332  
 RESULT 5  
 ID 044382 PRELIMINARY; PRT; 510 AA.  
 AC 044382;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE SLMB  
 GN SLMB OR SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly)  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96121115; PubMed=9461217;  
 RA Jiang J., Struhl G.;  
 RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-  
 box/WD40-repeat protein Slmb.";  
 RT Nature 391:493-496(1998).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF032878; AAC38855.1; .  
 DR Flybase: FBgn0023423; slmb.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001880; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat: WD repeat.  
 SQ SEQUENCE 510 AA; 58952 MW; DBB0243D3730A5E8 CRC64;

Query Match 90.2%; Score 138; DB 5; Length 510;  
 Best Local Similarity 89.3%; Pred. No. 6,9e-13;  
 Matches 25; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 GHTGSLCQYDERVITITSSSDSTVRW 29  
 Db 248 GHTGSLCQYDERVITITSSSDSTVRW 275  
 RESULT 6  
 ID 09VDE3 PRELIMINARY; PRT; 510 AA.  
 AC 09VDE3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE SLMB protein (SLMB).  
 GN SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton R.C., Rogers Y.-H.C., Blazer R.G., Chape M., Pfeiffer B.D.,  
 RA Brandon R.C., Rogers Y.-H.C., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Wan K.H., Doyle C., Baxter E.G., Andrews-Plankoch C., Baldwin D.,  
 RA Avril J.F., Aghayani A., An H.-J., Andrews-Plankoch C., Beasley E.M.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Bolshakov S.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,  
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,  
 RA Palazon M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svltkas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Ovary, AND IMAGINAL DISC;  
 RX MEDLINE=20245299; PubMed=10781936;  
 RA Mielich I., Limbong-Bouchon B.;  
 RT "Drosophila null slmb clones transiently deregulate Hedgehog-  
 independent transcription of all limb discs, and induce

RT decapentaplegic transcription linked to imaginal disc regeneration."  
 CC Mech Dev. 93:15-26(2000).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AE003733; AAF55853.1; -  
 DR EMBL; AF222924; AAF63213.1; -  
 DR EMBL; AF222923; AAF63213.1; -  
 DR FlyBase; FBgn0023423; slmb.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00646; F-box; 1.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 510 AA; 58952 MW; F4D5DF126F58A012 CRC64;

Query Match  
 Best Local Similarity 89.3%; Score 138; DB 5; Length 510;  
 Matches 25; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GHTGSVLCLOYDERVITIGSSDSSTVRW 29  
 DB 248 GHTGSVLCLOYDDKVIITSSSDSTVRW 275

RESULT 7  
 ID 080UN4 PRELIMINARY; PRT; 38 AA.  
 AC 080UN4;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Beta-TRCP protein (Fragment).  
 GN BETA-TRCP  
 OS Xenopus laevis (African clawed frog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 CC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Carnevali F.;  
 RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ballarino M.;  
 RA "Analisi strutturale e funzionale del gene beta-TRCP in Xenopus  
 laevis".  
 RT Department of Genetics and Molecular Biology "Charles Darwin",  
 RT University of Rome, Rome, Italy.  
 RU EMBL; A428937; CAD21934.1; -  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 1.  
 DR ProDom; PD000018; WD40; 1.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 FT NON\_TER 1  
 FT NON\_TER 38  
 SQ SEQUENCE 38 AA; 4270 MW; F825C0B481D54F03 CRC64;

Query Match  
 Best Local Similarity 81.0%; Score 124; DB 13; Length 38;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GHTGSVLCLOYDERVITIGSSDSST 25  
 DB 15 GHTGSVLCLOYDERVITIGSSDSST 38

RESULT 8  
 ID 09B354 PRELIMINARY; PRT; 430 AA.  
 AC 09B354;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Beta transducin repeat-containing protein (Fragment).  
 OS Heterodera glycyines (Soybean cyst nematode).  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;  
 CC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.  
 OC NCBI\_TaxID=51029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kovaleva E.S.; Yakovlev A.G.; Masler E.P.;  
 RT "Plant parasitic nematode b-TRCP."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AF339101; AAK26376.1; -  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER 1  
 FT NON\_TER 430  
 SQ SEQUENCE 430 AA; 47916 MW; 4ABC3F2DFE3A50B CRC64;

Query Match  
 Best Local Similarity 81.0%; Score 124; DB 5; Length 430;  
 Matches 21; Conservative 72.4%; Pred. No. 8.6e-11;  
 Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TGHGVSVCLOYDERVITIGSSDSSTVRW 29  
 DB 128 SGHTGSVLCLOYDNOMITAGSSDSSTVRW 156

RESULT 9  
 ID 0922C7 PRELIMINARY; PRT; 252 AA.  
 AC 0922C7;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Similar to f-box and WD-40 domain protein 1B (Fragment).  
 GN FXW1B.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; BC008552; AA08552.1; -  
 DR MGD; MGI:2144023; Fbxw1b.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 6.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR ProDom; PD000018; WD40; 3.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER 1  
 FT NON\_TER 252  
 SQ SEQUENCE 252 AA; 28424 MW; F71737CB07A9F75F CRC64;



Query Match  
Best Local Similarity 50.0%; Score 97; DB 11; Length 252;  
Matches 11; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 2 GHTGSVLCQYDERVITIGSSDSTVRW 29  
DB 111 GHKRGICACGYRDLVVGSSSDNTIRLW 138

RESULT 10  
Q8ST25 PRELIMINARY; PRT; 280 AA.

AC Q8ST25; 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
GN EC011\_1970.  
OS Encephalitozoon cuniculi.  
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.

OC NCBI\_TaxID=6035;  
RN [1] SEQUENCE FROM N.A.  
RP STRAIN=GB-M1;  
RA Genoscope;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2] SEQUENCE FROM N.A.  
RP STRAIN=GB-M1;  
RA MEDLINE-21576510; PubMed-11719806;  
RA Kacinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
RA Prenster G., Barbe V., Peyretailade E., Brotlier P., Wincker P.,  
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
RA Weissbach J., Vivares C.P.;  
RT "Genome sequence and gene compaction of the eukaryote parasite  
RT Encephalitozoon cuniculi.";  
RL Nature 414:450-453(2001).  
DR EMBL; AL590450; CAD26107.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 280 AA; 3139 MW; 50176DP5970F538 CRC64;

Query Match  
Best Local Similarity 58.2%; Score 89; DB 5; Length 280;  
Matches 17; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GHTGSVLCQYDERVITIGSSDSTVRW 29  
DB 94 GHMSSVLCQYDRIRIFSSSDCTIRW 121

RESULT 11  
Q8SW59 PRELIMINARY; PRT; 414 AA.

AC Q8SW59; 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
GN EC003\_0430.  
OS Encephalitozoon cuniculi.  
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.

OC NCBI\_TaxID=6035;  
RN [1] SEQUENCE FROM N.A.  
RP STRAIN=GB-M1;  
RA Genoscope;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2] SEQUENCE FROM N.A.  
RP STRAIN=GB-M1;  
RA MEDLINE-21576510; PubMed-11719806;  
RA Kacinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
RA Prenster G., Barbe V., Peyretailade E., Brotlier P., Wincker P.,

RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
RA Weissbach J., Vivares C.P.;  
RT "Genome sequence and gene compaction of the eukaryote parasite  
RT Encephalitozoon cuniculi.";  
RL Nature 414:450-453(2001).  
DR EMBL; AL590453; CAD26189.1; -  
SQ SEQUENCE 414 AA; 46065 MW; EBBCA38393981871 CRC64;

Query Match  
Best Local Similarity 57.2%; Score 87.5; DB 5; Length 414;  
Matches 17; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 TGHGTGSVLCQYDERVITIGSSDSTVRW 29  
DB 111 SGHEGPLYCVRG-GVITVGGSDSTVRFW 138

RESULT 12  
Q9NOX6 PRELIMINARY; PRT; 553 AA.

AC Q9NOX6; 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE CDNA FLJ11071 f1s, clone PLACE1004937, moderately similar to SEL-10  
DE protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI\_TaxID=9606;  
RN [1] SEQUENCE FROM N.A.  
RP TISSUE=PLACENTA;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Mitsuoka H., Ishii S., Sugano S., Aotsuka S., Yoshikawa Y.,  
RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.,  
RA "MDO human cDNA sequencing project.";  
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

CC -3- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
DR EMBL; AK001933; BAA91986.1; -  
DR InterPro: IPR001810; F-box.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00646; F-box; 1.  
DR Pfam: PF00400; WD40; 7.  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR PRODOM: PD000018; WD40; 2.  
DR SMART: SM00256; FBOX; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS50181; FBOX; 1.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.

QY 2 GHTGSVLCQYDERVITIGSSDSTVRW 29  
DB 305 GHTSVRCMHLERKRVVSGSRDATTIRW 332

Query Match  
Best Local Similarity 54.9%; Score 84; DB 4; Length 553;  
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHTGSVLCQYDERVITIGSSDSTVRW 29  
DB 305 GHTSVRCMHLERKRVVSGSRDATTIRW 332

RESULT 13  
Q96R12 PRELIMINARY; PRT; 561 AA.

AC Q96R12; 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE F-box protein FBX30 (Fragment).

```

Oy      2 GHTGSVLCLOQDERVIITGSSSDSTRVW 29
        |||::||: |: ::|| |:|||
Best Local Similarity 46.4%; Pred.No.0.0002; Mismatches 7; Indels 0; Gaps 0
Matches 13; Conservative 8;

Db      341 GHTSTVRCHMLHEKRVSGSRDATLRW 368

RESULT 15
096A16
ID      096A16          PRELIMINARY; PRT; 627 AA.
DC      096A16;
DT      01-DEC-2001 (TRMBLrel. 19, Created)
DT      01-DEC-2001 (TRMBLrel. 19, Last sequence update)
DE      01-JUN-2002 (TRMBLrel. 21, Last annotation update)
DR      Archipelago beta form (F-box protein FBW7).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RX      NCBI_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=21449047; PubMed=11565033;
RT      "Archipelago regulates Cyclin E levels in Drosophila and is mutated in
RL      human cancer cell lines."
RN      Nature 413:311-316(2001).
[2]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=20003061; PubMed=10531037;
RT      Winston J.T., Koeppe D.M., Zhu C., Elledge S.J., Harper J.W.;
RN      "A family of mammalian F-box proteins."
RN      Curr. Biol. 9:1180-1182(1999).
[3]
RP      SEQUENCE FROM N.A.
RA      Koeppe D.M., Winston J.T., Harper W., Elledge S.J.;
RN      Submitted (Apr-2001) to the EMBL/Genbank/DDbj databases.
CC      -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR      EMBL; AF411972; AAL06291.1; -.
DR      EMBL; AY033553; AAK57547.1; -.
DR      InterPro; IPR001810; F-box.
DR      InterPro; IPR001680; WD40.
DR      Pfam; PF00646; F-box; 1.
DR      Pfam; PF00400; WD40; 7.
DR      ProDom; PD00018; WD40; 2.
DR      PROSITE; PS50181; FBOX; 1.
DR      PROSITE; PS00678; WD_REPEATS_1; UNKNOWN-5.
DR      PROSITE; PS50082; WD_REPEATS_2; 7.
DR      PROSITE; PS50294; WD_REPEATS_REGION; 1.
KM      Repeat; WD repeat.
SQ      SEQUENCE 627 AA; 70324 MW; 3D4107C053381BED CRC64;

Query Match      54.9%; Score 84; DB 4; Length 627;
Best Local Similarity 46.4%; Pred.No.0.00022;
Matches 13; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Oy      2 GHTGSVLCLOQDERVIITGSSSDSTRVW 29
        |||::||: |: ::|| |:|||
Db      379 GHTSTVRCHMLHEKRVSGSRDATLRW 406

Search completed: April 11, 2003, 11:51:34
Job time : 8.85262 secs

```

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:20:44 ; Search time 9.55298 Seconds  
(without alignments)  
432.406 Million cell updates/sec

```
Title: US-09-601-168B-2_COPY_343_373
Perfect score: 172
Sequence: 1 LHHCEAVLHFRNNGMVTCSKDRSIAVMD 31
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

A.GeneSeq\_101002:\*

1: /SIDS2/gcgdata/genseq/genseq-emb1/AA1980.DAT.\*

2: /SIDS2/gcgdata/genseq/genseq-emb1/AA1981.DAT.\*

3: /SIDS2/gcgdata/genseq/genseq-emb1/AA1982.DAT.\*

4: /SIDS2/gcgdata/genseq/genseq-emb1/AA1983.DAT.\*

5: /SIDS2/gcgdata/genseq/genseq-emb1/AA1984.DAT.\*

6: /SIDS2/gcgdata/genseq/genseq-emb1/AA1985.DAT.\*

7: /SIDS2/gcgdata/genseq/genseq-emb1/AA1986.DAT.\*

8: /SIDS2/gcgdata/genseq/genseq-emb1/AA1987.DAT.\*

9: /SIDS2/gcgdata/genseq/genseq-emb1/AA1988.DAT.\*

10: /SIDS2/gcgdata/genseq/genseq-emb1/AA1989.DAT.\*

11: /SIDS2/gcgdata/genseq/genseq-emb1/AA1990.DAT.\*

12: /SIDS2/gcgdata/genseq/genseq-emb1/AA1991.DAT.\*

13: /SIDS2/gcgdata/genseq/genseq-emb1/AA1992.DAT.\*

14: /SIDS2/gcgdata/genseq/genseq-emb1/AA1993.DAT.\*

15: /SIDS2/gcgdata/genseq/genseq-emb1/AA1994.DAT.\*

16: /SIDS2/gcgdata/genseq/genseq-emb1/AA1995.DAT.\*

17: /SIDS2/gcgdata/genseq/genseq-emb1/AA1996.DAT.\*

18: /SIDS2/gcgdata/genseq/genseq-emb1/AA1997.DAT.\*

19: /SIDS2/gcgdata/genseq/genseq-emb1/AA1998.DAT.\*

20: /SIDS2/gcgdata/genseq/genseq-emb1/AA2000.DAT.\*

21: /SIDS2/gcgdata/genseq/genseq-emb1/AA2001.DAT.\*

22: /SIDS2/gcgdata/genseq/genseq-emb1/AA2002.DAT.\*

23: /SIDS2/gcgdata/genseq/genseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	172	100.0	510	22	ABR59857	Drosophila melanog
2	172	100.0	517	16	AAH85852	WD-40 domain-contig
3	172	100.0	559	20	AAW24054	A human beta- <i>cr</i> ans
4	172	100.0	569	21	AAH12812	Mouse ubiquitin li
5	172	100.0	569	21	AAH12813	Human beta-transduc
6	172	100.0	569	21	AAV96697	Human beta- <i>TRCP</i>
7	172	100.0	569	21	AAV83041	F-box protein FBP-
8	172	100.0	559	21	AAV83250	F-box protein hbetF
9	172	100.0	569	21	AAV83254	F-box protein FWD1
10	172	100.0	569	21	AAV44249	Human cell signal

11	177	100.0	569	22	AAH8289	Human Zf11 protein
12	172	100.0	579	22	AAH78583	Human protein SEQ
13	172	100.0	590	22	AAH00847	Human bone marrow
14	172	100.0	605	22	AAH78582	Human protein SEQ
15	172	100.0	608	22	AAH00960	Human bone marrow
16	172	100.0	632	22	AAH78584	Human protein SEQ
17	172	100.0	654	22	AAH79566	Human protein SEQ
18	172	100.0	654	22	AAH79567	Human protein SEQ
19	172	100.0	654	22	AAH79568	Peptide r11 from
20	168	97.7	30	16	AAH84937	Human E3 ubiqitin
21	152	88.4	542	22	AAH96666	Human protein SEQ
22	152	88.4	542	22	AAH79127	Human polypeptide
23	152	88.4	542	22	AAH40268	Human polypeptide
24	152	88.4	550	22	AAH41994	Human polypeptide
25	71	41.3	84	23	ABH32299	Human transcriptio
26	68	39.5	627	22	ABH64267	Drosophila melanog
27	66.5	38.7	412	22	AAH93214	Human polypeptide,
28	65	38.4	436	22	ABH59186	Drosophila melanog
29	65	37.8	640	21	AAH83252	F-box protein Mel3
30	64	37.2	110	23	ABH06489	Human ORF protein
31	64	37.2	357	22	AAH93201	Human polypeptide,
32	64	37.2	375	22	AAE02854	Human GTPase-assoc
33	64	37.2	566	22	AAH40285	Human polypeptide
34	63.5	36.9	459	22	AAH82973	Human G protein as
35	63.5	36.9	459	22	AAH93316	Human protein sequ
36	63.5	36.9	459	22	AAH93830	Human protein sequ
37	63.5	36.9	459	22	AAH94812	Human protein sequ
38	63	36.6	117	21	AAH02191	Human secreted pro
39	63	36.6	199	21	AAH07146	Arabidopsis thalia
40	63	36.6	204	21	AAH07145	Arabidopsis thalia
41	63	36.6	299	21	AAH51154	Arabidopsis thalia
42	63	36.6	304	21	AAH51153	Arabidopsis thalia
43	63	36.6	330	21	AAH51148	Arabidopsis thalia
44	62.5	36.3	1326	22	ABH67237	Drosophila melanog
45	62.5	36.3	1326	22	ABH67238	Drosophila melanog

## ALIGNMENTS

PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
XX	
XX	
DR	N-PSDB: ABL03960.
DR	WPI; 2001-656860/75.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EM;
PA	(PEKE ) PE CORP NY.
XX	
PR	11-JUL-2000; 2000US-0614150.
PR	23-MAR-2000; 2000US-191637P.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PD	27-SEP-2001.
XX	
PN	WO200171042-A2.
OS	Drosophila melanogaster.
XX	
XX	
KM	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.
DE	
XX	
DT	26-MAR-2002 (first entry)
XX	
ID	ABB59857 standard; Protein; 510 AA.
AC	ABB59857;
XX	
XX	
ABBS59857	
RESULT 1	

PT interactions -  
 XX  
 PS Disclosure: SEQ ID NO 6363; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (AB101840-AB116175).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPD  
 CC at ftp.wipd.int/pub/published\_pct\_sequences.  
 XX  
 SO Sequence 510 AA;  
 Query Match 100.0%; Score 172; DB 22; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LIHCEAVLHFRNGMVTCSKDRSLAWD 31  
 Db 286 LIHCEAVLHFRNGMVTCSKDRSLAWD 316  
 RESULT 2  
 AAR85852  
 ID AAR85852 standard; peptide; 517 AA.  
 XX  
 AC AAR85852:  
 XX  
 DT 13-SEP-1996 (first entry)  
 XX  
 DE WD-40 domain-contg. beta-TRCP protein.  
 XX  
 KW WD40 repeat region; beta-transducin; protein-protein interaction; drug;  
 KW intracellular signalling; protein kinase C; homology; motif; modulator;  
 KW receptors of activated protein kinase; enzyme activity; isozyme; human.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9521252-A2.  
 XX  
 PD 10-AUG-1995.  
 XX  
 PF 31-JAN-1995; 95WO-US01210.  
 XX  
 PR 01-FEB-1994; 94US-0190802.  
 XX  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 XX  
 PI Mochly-Rosen D, Ron D;  
 XX  
 DR WPI: 1995-283772/37.  
 XX  
 PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the  
 PT activity of a protein, eg. protein kinase C, which interacts with a  
 PT protein contg. a WD-40 region.  
 XX  
 PS Example 5; Page 80-82; 351pp; English.  
 XX  
 CC Proteins AAR85851-92 are protein which contain at least one WD-40 (also  
 CC called beta-transducin homologous) amino acid repeat motifs. The WD-40  
 CC regions are involved in protein-protein interactions between proteins  
 CC involved in intracellular signalling. An example of such an interaction  
 CC is between protein kinase C and receptors of activated protein kinase  
 CC (RACK), esp. RACK-1 (AAR85850). Proteins AAR85851-82 were isolated based  
 CC on homology with beta-transducin, whereas proteins AAR85882-92 were  
 CC isolated based on homology with the WD-40 consensus sequence (AAR85893).  
 CC The proteins were used to construct the peptides AAR84928-R85063 and  
 CC AAR85786-R85842. The peptides can be used to identify target proteins

CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of  
 CC proteins involved in protein-protein interaction and to screen for drugs  
 CC that will affect protein-protein interaction involving WD-40 domains.  
 XX  
 SO Sequence 517 AA;  
 Query Match 100.0%; Score 172; DB 16; Length 517;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LIHCEAVLHFRNGMVTCSKDRSLAWD 31  
 Db 307 LIHCEAVLHFRNGMVTCSKDRSLAWD 337  
 RESULT 3  
 AAY24054  
 ID AAY24054 standard; Protein; 569 AA.  
 XX  
 AC AAY24054:  
 XX  
 DT 30-SEP-1999 (first entry)  
 XX  
 DE A human beta-transducin repeat containing protein.  
 XX  
 KW Beta-transducin repeat containing protein; beta-Trcp; SKP1p;  
 KW proteasome degradation pathway; Vpu protein; beta-catenin;  
 KW human immune deficiency virus-1; HIV-1; cellular protein; IKAPPAB;  
 KW ubiquitinylation; phosphorylated protein; tumour; apoptosis; Alzheimer's;  
 KW antiviral; antitumour; cell cycle regulation; protein degradation;  
 KW and anti-inflammatory; osteo-articular inflammation; acute inflammation;  
 KW tumour necrosis factor.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Region Location/Qualifiers  
 FT 147..191  
 FT /note= "F box sequence"  
 FT 259..292  
 FT /note= "WD motif"  
 FT 304..332  
 FT /note= "WD motif"  
 FT 343..372  
 FT /note= "WD motif"  
 FT 387..415  
 FT /note= "WD motif"  
 FT 427..455  
 FT /note= "WD motif"  
 FT 467..492  
 FT /note= "WD motif"  
 FT 516..544  
 FT /note= "WD motif"  
 FT Region  
 XX  
 PN WO9938969-A1.  
 XX  
 PD 05-AUG-1999.  
 XX  
 PF 29-JAN-1999; 99WO-FR00196.  
 XX  
 PR 09-DEC-1998; 98FR-0015545.  
 XX  
 PR 30-JAN-1998; 98FR-0001100.  
 XX  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Arenzana Seisdedos F, Benarous R, Concordet J, Durand H;  
 PI Kroll M, Margottin F;  
 XX  
 DR WPI: 1999-469329/39.  
 DR N-PSDB: AAX86501.  
 XX  
 PT New human beta-transducin repeat containing protein and its  
 PT fragments useful as, or to screen for, antiviral, antitumour,

PT anti-inflammatory and anti-Alzheimer's agents  
XX  
PS Claim 1; Page 60-61; 71pp; French.  
XX  
CC The present sequence represents a human beta-transducin repeat containing  
CC protein (beta-TrCP). The protein directs proteins to the proteasome  
CC degradation pathways. The protein is able to interact with the Vpu  
CC protein of human immune deficiency virus-1 (HIV-1), cellular proteins  
CC IkappaB or beta-catenin (bc) and/or protein Skp1. The protein controls  
CC ubiquitinylation of phosphorylated proteins and thus their targeting to  
CC proteasomes for degradation. Depending on whether the process is  
CC inhibited or promoted, the result may be delayed breakdown of CD4 (in  
CC cases of HIV-1 infection); increased activity of Ikb (and thus reduced  
CC activity of NFkappaB) and increased degradation of mutant bc in tumour  
CC cells, or increased bc survival (and reduced apoptosis) in Alzheimer's  
CC patients. The beta-TrCP protein, and its active peptide fragments, or its  
CC nucleic acid, are used to screen for anti HIV-1 agents (antivirals),  
CC antitumour agents that disrupt cell cycle regulation or protein  
CC degradation in human tumour cells, and anti-inflammatory agents that  
CC disrupt activation by NFkappaB. Fragments of the protein are also  
CC useful for treating osteo-articular inflammation or acute inflammation  
CC associated with release of tumour necrosis factor.  
XX  
SQ Sequence 569 AA:  
Query Match 100.0%; Score 172; DB 20; Length 569;  
Best Local Similarity 100.0%; Pred. No. 5.4e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LIHHCeAVLHLRFNNGMWTCSKDRSIAVWD 31  
DB 343 LIHHCeAVLHLRFNNGMWTCSKDRSIAVWD 373  
RESULT 4  
AAB12812  
ID AAB12812 standard; protein; 569 AA.  
XX  
AC AAB12812;  
XX  
DT 27-NOV-2000 (first entry)  
XX  
DE Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.  
XX  
XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;  
XX beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;  
XX gene therapy; colon cancer; beta-transducin repeat containing protein;  
XX beta-TrCP.  
XX  
OS Mus musculus.  
XX  
PN JP2000166542-A.  
XX  
PD 20-JUN-2000.  
XX  
PF 02-DEC-1998; 98JP-0343437.  
XX  
PR 02-DEC-1998; 98JP-0343437.  
XX  
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
XX  
XX WPI: 2000-485550/43.  
XX  
DR N-PSDB; AAA73131.  
XX  
PT F-box protein of ubiquitin ligase SCF complex which promotes the  
XX ubiquitination of IkappaB or beta-catenin  
XX  
PS Claim 2; Page 9-10; 19pp; Japanese.  
XX  
CC The present invention describes an F-box motif protein of ubiquitin  
XX ligase SCF complex which promotes the ubiquitination of IkappaB or  
XX beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
XX complex (SCF complex) of F-box protein containing F-box motif and WD40

CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
CC the gene therapy of colon cancer by being recombined to a virus vector.  
XX  
SQ Sequence 569 AA:  
Query Match 100.0%; Score 172; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 5.4e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LIHHCeAVLHLRFNNGMWTCSKDRSIAVWD 31  
DB 343 LIHHCeAVLHLRFNNGMWTCSKDRSIAVWD 373  
RESULT 5  
AAB12813  
ID AAB12813 standard; protein; 569 AA.  
XX  
AC AAB12813;  
XX  
DT 27-NOV-2000 (first entry)  
XX  
DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.  
XX  
XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;  
XX beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;  
XX gene therapy; colon cancer; beta-transducin repeat containing protein;  
XX beta-TrCP.  
XX  
OS Homo sapiens.  
XX  
PN JP2000166542-A.  
XX  
PD 20-JUN-2000.  
XX  
PF 02-DEC-1998; 98JP-0343437.  
XX  
PR 02-DEC-1998; 98JP-0343437.  
XX  
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
XX  
XX WPI: 2000-485550/43.  
XX  
DR N-PSDB; AAA73132.  
XX  
PT F-box protein of ubiquitin ligase SCF complex which promotes the  
XX ubiquitination of IkappaB or beta-catenin  
XX  
PS Claim 3; Page 10-12; 19pp; Japanese.  
XX  
XX The present invention describes an F-box motif protein of ubiquitin  
XX ligase SCF complex which promotes the ubiquitination of IkappaB or  
XX beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
XX complex (SCF complex) of F-box protein containing F-box motif and WD40  
XX repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
XX or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
XX ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
XX repeat containing protein (beta-TrCP)). The F-box protein can be used for  
XX the gene therapy of colon cancer by being recombined to a virus vector.  
XX  
SQ Sequence 569 AA:  
Query Match 100.0%; Score 172; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 5.4e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LIHHCeAVLHLRFNNGMWTCSKDRSIAVWD 31  
DB 343 LIHHCeAVLHLRFNNGMWTCSKDRSIAVWD 373



OS Homo sapiens.  
 XX  
 PN WO200022110-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 XX 08-OCT-1999; 99WO-US23705.  
 XX  
 PR 09-OCT-1998; 98US-0103787.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Zhou P, Howley P;  
 XX  
 DR WPI: 2000-317970/27.  
 DR N-PSDB; AA293710.  
 XX  
 PT Targeting degradation of polypeptide useful for treating cancer and  
 PT other proliferative disorders, involves conjugating polypeptide with  
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic  
 PT compound  
 XX  
 PS Claim 9; Page 171; 185pp; English.  
 XX  
 CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin  
 CC ligases) which can be used for the targeted degradation of a target  
 CC polypeptide in vivo. Targeted degradation is achieved by expressing  
 CC the ubiquitin ligase in a cell linked to the interaction domain of  
 CC the target polypeptide and thereby recruiting the target polypeptide  
 CC to the ubiquitin ligase. Such methods are useful for decreasing or  
 CC increasing the level of a target polypeptide and for creating and  
 CC expressing a destabilized polypeptide which is subjected to SCF  
 CC mediated proteolysis. Degrading any desired protein in a cell is  
 CC useful for preventing or treating diseases caused by the presence of  
 CC abnormal amount of the specific polypeptides, for drug discovery and  
 CC for gene therapy. Diseases treated include cancer, by degradation of  
 CC oncoproteins, Huntington's disease, other proliferative disorders and  
 CC microbial infections. The method provides a quick and easy  
 CC alternative to gene knockout technology. The target polypeptide can  
 CC be degraded at all stages, or a specific stage, of development in the  
 CC mature animal.  
 CC  
 XX  
 SQ Sequence 569 AA;  
 XX  
 Query Match 100.0%; Score 172; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHCEAVLHLRFNNGMWTCKDRSIAVWD 31  
 DB 343 LIHCEAVLHLRFNNGMWTCKDRSIAVWD 373  
 |||||||||||||||||||||||||||||

RESULT 9  
 ID AAY83254 standard; Protein; 569 AA.  
 XX  
 AC AAY83254;  
 XX  
 DT 16-AUG-2000 (first entry)  
 XX  
 DE F-box protein FWD1p.  
 XX  
 XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;  
 KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;  
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;  
 KW mouse; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200022110-A2.  
 XX  
 PD 20-APR-2000.

XX  
 PF 08-OCT-1999; 99WO-US23705.  
 XX  
 PR 09-OCT-1998; 98US-0103787.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Zhou P, Howley P;  
 XX  
 DR WPI: 2000-317970/27.  
 DR N-PSDB; AA293714.  
 XX  
 PT Targeting degradation of polypeptide useful for treating cancer and  
 PT other proliferative disorders, involves conjugating polypeptide with  
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic  
 PT compound  
 XX  
 PS Claim 9; Page 184-185; 185pp; English.  
 XX  
 CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin  
 CC ligases) which can be used for the targeted degradation of a target  
 CC polypeptide in vivo. Targeted degradation is achieved by expressing  
 CC the ubiquitin ligase in a cell linked to the interaction domain of  
 CC the target polypeptide and thereby recruiting the target polypeptide  
 CC to the ubiquitin ligase. Such methods are useful for decreasing or  
 CC increasing the level of a target polypeptide and for creating and  
 CC expressing a destabilized polypeptide which is subjected to SCF  
 CC mediated proteolysis. Degrading any desired protein in a cell is  
 CC useful for preventing or treating diseases caused by the presence of  
 CC abnormal amount of the specific polypeptides, for drug discovery and  
 CC for gene therapy. Diseases treated include cancer, by degradation of  
 CC oncoproteins, Huntington's disease, other proliferative disorders and  
 CC microbial infections. The method provides a quick and easy  
 CC alternative to gene knockout technology. The target polypeptide can  
 CC be degraded at all stages, or a specific stage, of development in the  
 CC mature animal.  
 CC  
 XX  
 SQ Sequence 569 AA;  
 XX  
 Query Match 100.0%; Score 172; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHCEAVLHLRFNNGMWTCKDRSIAVWD 31  
 DB 343 LIHCEAVLHLRFNNGMWTCKDRSIAVWD 373  
 |||||||||||||||||||||||||||||

RESULT 10  
 ID AAY44249 standard; Protein; 569 AA.  
 XX  
 AC AAY44249;  
 XX  
 DT 28-FEB-2000 (first entry)  
 XX  
 DE Human cell signalling protein-12.  
 XX  
 XX Cell signalling protein-12; CSIGP-12; cell proliferation;  
 KW inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;  
 KW arteriosclerosis; Addison's disease; multiple sclerosis.  
 KW  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 19  
 FT Modified-site 39 /note= "Potential phosphorylation site"  
 FT Modified-site 91 /note= "Potential phosphorylation site"  
 FT Modified-site 109 /note= "Potential phosphorylation site"  
 FT Modified-site 109 /note= "Potential phosphorylation site"

FT Modified-site 162 /note= "Potential phosphorylation site"  
 FT Modified-site 266 /note= "Potential phosphorylation site"  
 FT Modified-site 288 /note= "Potential phosphorylation site"  
 FT Modified-site 328 /note= "Potential phosphorylation site"  
 FT Modified-site 376 /note= "Potential phosphorylation site"  
 FT Modified-site 381 /note= "Potential phosphorylation site"  
 FT Modified-site 411 /note= "Potential phosphorylation site"  
 FT Modified-site 418 /note= "Potential phosphorylation site"  
 FT Modified-site 451 /note= "Potential phosphorylation site"  
 FT Modified-site 514 /note= "Potential phosphorylation site"  
 FT Modified-site 519 /note= "Potential phosphorylation site"  
 FT Modified-site 535 /note= "Potential phosphorylation site"  
 FT Modified-site 536 /note= "Potential phosphorylation site"  
 FT Modified-site 17 /note= "Potential phosphorylation site"  
 FT Modified-site 77 /note= "Potential glycosylation site"  
 FT Modified-site 416 /note= "Potential glycosylation site"  
 FT Modified-site 320...334 /note= "Potential glycosylation site"  
 FT Region /label= "Signature-sequence"  
 FT Region 360...374 /label= "Signature-sequence"  
 FT Region 403...417 /label= "Signature-sequence"  
 FT Region 443...457 /label= "Signature-sequence"  
 FT Region 483...497 /label= "Signature-sequence"  
 FT Region /label= "Signature-sequence"  
 FT Region 532...546 /label= "Signature-sequence"  
 FT W0958558-A2.  
 PN 18-NOV-1999.  
 XX 13-MAY-1999; 99WO-US10567.  
 XX 13-MAY-1998; 98US-0085343.  
 PR 26-AUG-1998; 98US-0098010.  
 XX (INCY-) INCYTE PHARM INC.  
 XX Bandman O, Hillman JL, Lai P, Yue H, Tang YT, Patterson C;  
 PI Baughman MR, Yang J;  
 DR WPI: 2000-086432/07.  
 DR N-PSDB; AAZ29233.  
 XX Human cell signaling proteins useful for, e.g. diagnosing cell  
 PT proliferative and inflammatory disorders -  
 XX Claim 1: Page 77-78; 90pp; English.  
 XX The present sequence is cell signalling protein-12 (CSIGP-12) encoded  
 CC by cDNA obtained from Incyte clone 3239149 of COAUC01 library. It is  
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is  
 CC found to be homologous to beta-transducin repeats containing  
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as  
 CC hybridisation probe for detecting CSIGP related sequences or allelic

CC variants. Recombinant CSIGP can be produced in host cells by transforming  
 CC them with genetically engineered vectors. Agonists or antagonists can be  
 CC used in the treatment of cell proliferative and inflammatory disorders  
 CC associated with decreased or increased CSIGP expression. CSIGP is used in  
 CC the diagnosis, prevention and treatment of cell proliferative disorders  
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.  
 XX  
 SQ Sequence 569 AA;  
 Query Match 100.0%; Score 172; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 5,4e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LIHCEAVLHLRFNGMWTGSKDRSIAWVD 31  
 DB 343 LIHCEAVLHLRFNGMWTGSKDRSIAWVD 373  
 RESULT 11  
 AAB48298  
 ID AAB48298 standard; protein; 569 AA.  
 XX AAB48298;  
 AC  
 XX 02-APR-2001 (first entry)  
 DE Human ZP11 protein.  
 XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; 2F;  
 KW CUL-1; Cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
 KW Bad; Bcl-2; tumour; cytostatic.  
 XX Homo sapiens.  
 OS  
 XX WO200075184-A1.  
 PN 14-DEC-2000.  
 PD 05-JUN-2000; 2000WO-US15449.  
 XX 04-JUN-1999; 99US-0137494.  
 PR (UYA) UNITV YALE.  
 XX Zhang H, Tsvetkov LM, Kondo T;  
 PI WPI: 2001-061703/07.  
 DR N-PSDB; AAC84610.  
 XX Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
 PT involves altering levels of proteins such as S-phase kinase associated  
 PT proteins 1, 2 and cullin/CDC53 proteins -  
 XX Claim 3; Page 130-132; 162pp; English.  
 XX The invention relates to methods of altering the polypeptide levels in a  
 CC cell, using proteins selected from S-phase kinase associated proteins 1  
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
 CC cullin/CDC53 family of proteins). The method is useful for altering the  
 CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for  
 CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
 CC that modulate interactions between SKP and target proteins are useful for  
 CC treating tumours.  
 XX  
 SQ Sequence 569 AA;  
 Query Match 100.0%; Score 172; DB 22; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 5,4e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LIHCEAVLHLRFNGMWTGSKDRSIAWVD 31



Db 343 LIHHCeAVLHLRFNNGMWTGCSKDRSIAVMD 373

RESULT 12  
ID AAM78583  
AA: AAM78583 standard; Protein: 579 AA.

AC AAM78583;  
DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1245.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

PN WO200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US04098.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Gao Y, Ma Y;  
PI Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

DR WPI: 2001-476283/51.  
N-PSDB: AAK51716.

XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
PS Claim 20; Page 3504-3505; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.

XX Sequence 579 AA;

Query Match 100.0%; Score 172; DB 22; Length 579;  
Best local Similarity 100.0%; Pred. No. 5,6e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHHCeAVLHLRFNNGMWTGCSKDRSIAVMD 31  
|||||

Db 353 LIHHCeAVLHLRFNNGMWTGCSKDRSIAVMD 383

RESULT 13  
ID AAM00847  
AA: AAM00847 standard; Protein: 590 AA.

AC AAM00847;  
DT 01-OCT-2001 (first entry)

DE Human bone marrow protein, SEQ ID NO: 210.

XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;  
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;  
KW immunosuppressive; gene therapy; cytokine cell proliferation;  
KW cell differentiation modulator; immune disorder; infection; cancer;  
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.

OS Homo sapiens.

PN WO200153453-A2.

PD 26-JUL-2001.

PF 23-DEC-2000; 2000WO-US34960.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 30-NOV-2000; 2000US-0250583.

XX (HYSE-) HYSEQ INC.

XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao Q, Zhou P, Drmanac RT;

DR WPI: 2001-488707/53.  
N-PSDB: AAH89966.

XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
PT for treating e.g. cancer and immune deficiency disorders -  
PS Claim 10; Page 354-355; 648pp; English.

XX The present sequence is one of 251 novel human polypeptides encoded  
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the  
CC polypeptide encoded by it are useful in the treatment of various  
CC immune deficiencies and disorders. The deficiencies and disorders may  
CC be genetic, or may result from an autoimmune disorder, a coagulation  
CC infection, or may result from an autoimmune disorder, a coagulation  
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
CC suppression of an inflammatory response or treatment of a nervous  
CC system disorder such as Alzheimer's disease. Detection of the presence  
CC or increased expression of the polynucleotide or the protein it  
CC encodes is useful for the diagnosis and/or prognosis of one  
CC or more types of cancer. The polynucleotide and polypeptide can be  
CC used as nutritional sources or supplements and in the screening of  
CC chemical compounds as potential drugs.

XX Sequence 590 AA;

Query Match 100.0%; Score 172; DB 22; Length 590;  
Best local Similarity 100.0%; Pred. No. 5,7e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHHCeAVLHLRFNNGMWTGCSKDRSIAVMD 31  
|||||

Db 364 LIHCEAVLHLRFNNGMVTCSKDRSTAVMD 394

## RESULT 14

AA078582

ID AAM78582 standard; Protein; 605 AA.

XX AAM78582;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1244.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wehrman T, Goodrich R;

XX WPI: 2001-476283/51.

DR N-PSDB; AAK51715.

XX Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

XX Claim 20; Page 3503-3504; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM78323-AA080302) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

CC (AAM80020) are omitted as the relevant pages from the sequence listing

CC were missing at the time of publication.

XX Sequence 605 AA;

XX Query Match 100.0%; Score 172; DB 22; Length 605;

XX Best Local Similarity 100.0%; Pred. No. 5.9e-18;

XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 379 LIHCEAVLHLRFNNGMVTCSKDRSTAVMD 409

## RESULT 15

AA000960

ID AAM00960 standard; Protein; 608 AA.

XX AAM00960;

XX 01-OCT-2001 (first entry)

XX Human bone marrow protein, SEQ ID NO: 436.

XX Human: bone marrow; anti-inflammatory; cytostatic; neuroprotective;

KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;

KW immunosuppressive; gene therapy; cytokine cell proliferation;

KW cell differentiation modulator; immune disorder; infection; cancer;

KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.

XX Homo sapiens.

XX WO200153453-A2.

XX 26-JUL-2001.

XX 23-DEC-2000; 2000WO-US34960.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 30-NOV-2000; 2000US-0250583.

XX (HYSE-) HYSEQ INC.

XX Ford JE, Boyle BJ, Tang YF, Liu C, Asundi V, Chen R, Ma Y;

PI Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Drmanac RT;

XX WPI: 2001-488707/53.

DR N-PSDB; AAK90079.

XX Novel bone marrow-expressed polynucleotides and polypeptides, useful

PT for treating e.g. cancer and immune deficiency disorders -

XX Claim 10; Page 523-524; 648pp; English.

XX The present sequence is one of 251 novel human polypeptides encoded

CC by a bone marrow-expressed polynucleotide. The polynucleotide and the

CC polypeptide encoded by it are useful in the treatment of various

CC immune deficiencies and disorders. The deficiencies and disorders may

CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal

CC infection, or may result from an autoimmune disorder, a coagulation

CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,

CC suppression of an inflammatory response or treatment of a nervous

CC system disorder such as Alzheimer's disease. Detection of the presence

CC or increased expression of the polynucleotide or the protein it

CC encodes is useful for the diagnosis and/or prognosis of one

CC or more types of cancer. The polynucleotide and polypeptide can be

CC used as nutritional sources or supplements and in the screening of

CC chemical compounds as potential drugs.

XX Sequence 608 AA;

XX Query Match 100.0%; Score 172; DB 22; Length 608;

XX Best Local Similarity 100.0%; Pred. No. 5.9e-16;

XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 382 LIHCEAVLHLRFNNGMVTCSKDRSTAVMD 412

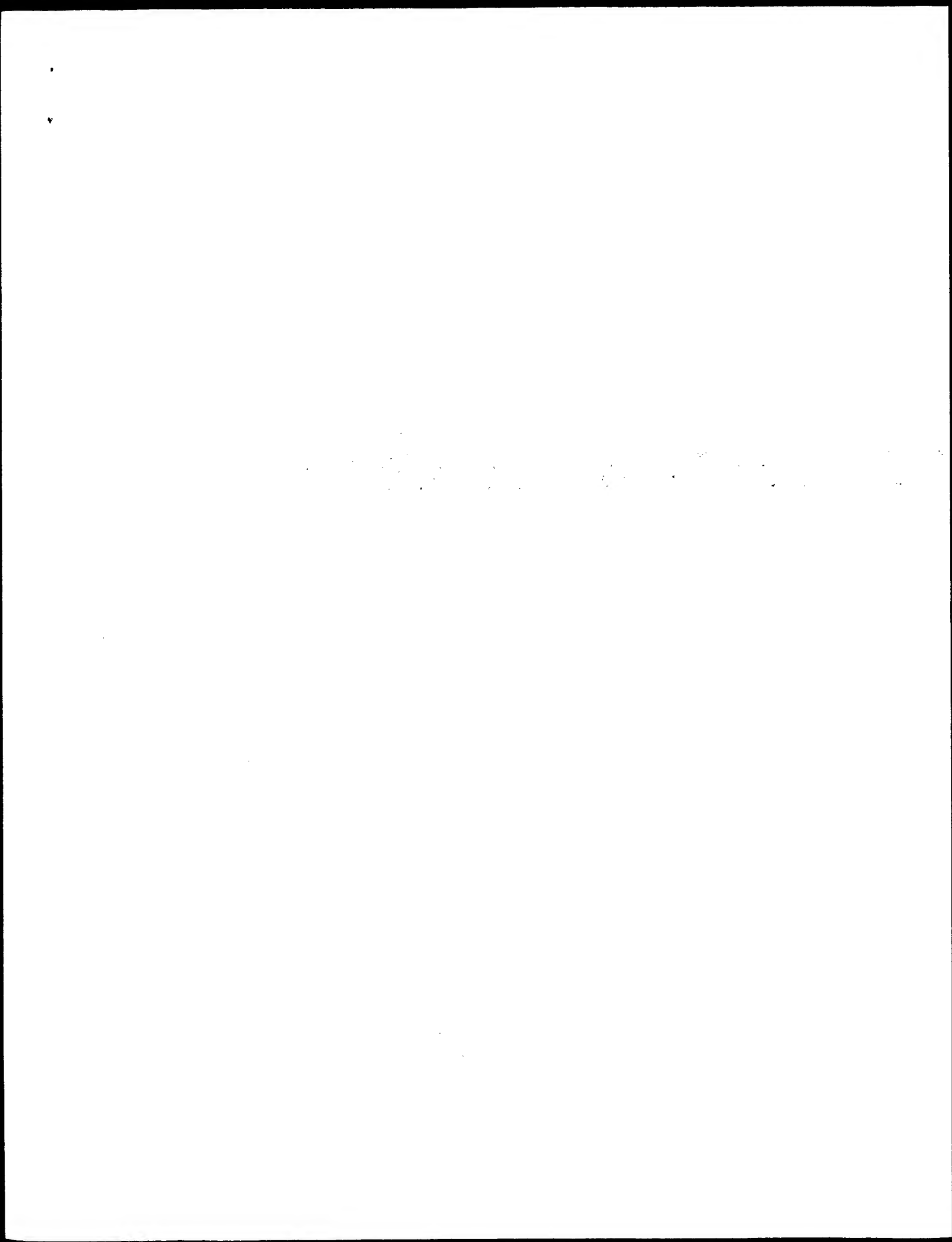
Fri Apr 11 13:31:09 2003

us-09-601-168b-2\_copy\_343\_373.rag

Page 9

Search completed: April 11, 2003, 11:48:15  
Job time : 10.553 secs

---



GenCore version 5.1.4-P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 3.2095 Seconds

(without alignments)  
284.191 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_343\_373

Perfect score: 172  
Sequence: 1 LHHCEAVLHLRFNNGMMVTCCKDRSLAVMD 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
7: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	100.0	517	1	US-08-190-802A-30
2	172	100.0	517	4	US-08-477-346-30
3	172	100.0	517	4	US-08-473-089-30
4	172	100.0	517	4	US-08-487-072A-30
5	168	97.7	30	1	US-08-190-802A-85
6	168	97.7	30	4	US-08-477-346-85
7	168	97.7	30	4	US-08-473-089-85
8	168	97.7	30	4	US-08-487-072A-85
9	65	37.8	640	4	US-09-177-165A-30
10	64	37.2	375	4	US-09-063-743-1
11	64	37.2	375	4	US-09-590-540-1
12	60	34.9	2629	2	US-08-751-189-4
13	60	34.9	2629	2	US-09-060-836-4
14	60	34.9	2629	4	US-09-184-445-4
15	58.5	34.0	421	4	US-09-302-769-14
16	57.5	33.4	37	1	US-08-190-802A-206
17	57.5	33.4	37	4	US-08-477-346-206
18	57.5	33.4	37	4	US-08-473-089-206
19	57.5	33.4	37	4	US-08-487-072A-206
20	57.5	33.4	576	1	US-08-190-802A-56
21	57.5	33.4	576	4	US-08-477-346-56
22	57.5	33.4	576	4	US-08-473-089-56
23	57.5	33.4	576	4	US-08-487-072A-56
24	57	33.1	431	1	US-08-190-802A-37
25	57	33.1	431	4	US-08-477-346-37
26	57	33.1	431	4	US-08-473-089-37
27	57	33.1	431	4	US-08-487-072A-37

28	57	33.1	1194	4	US-09-092-508-2	Sequence 2, Appl
29	57	33.1	1194	4	US-09-435-115-2	Sequence 2, Appl
30	57	33.1	1194	4	US-09-069-023-26	Sequence 26, Appl
31	57	33.1	1194	4	US-09-098-310-2	Sequence 2, Appl
32	57	33.1	1205	4	US-09-092-508-16	Sequence 16, Appl
33	57	33.1	1205	4	US-09-435-115-16	Sequence 16, Appl
34	56	32.6	31	1	US-08-190-802A-254	Sequence 254, App
35	56	32.6	31	4	US-08-477-346-254	Sequence 254, App
36	56	32.6	31	4	US-08-473-089-254	Sequence 254, App
37	56	32.6	31	4	US-08-487-072A-254	Sequence 52, Appl
38	56	32.6	422	1	US-08-477-346-52	Sequence 52, Appl
39	56	32.6	422	4	US-08-473-089-52	Sequence 52, Appl
40	56	32.6	422	4	US-08-487-072A-52	Sequence 52, Appl
41	56	32.6	422	4	US-08-190-802A-66	Sequence 66, Appl
42	56	32.6	514	1	US-08-477-346-66	Sequence 66, Appl
43	56	32.6	514	4	US-08-473-089-66	Sequence 66, Appl
44	56	32.6	514	4	US-08-487-072A-66	Sequence 66, Appl
45	56	32.6	514	4	US-08-487-072A-66	Sequence 66, Appl

#### ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
; Sequence 30, Application US/08190802A  
; Patent No. 5519003  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: P.O. Box 60850  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-0850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190, 802A  
; FILING DATE: 01-FEB-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fadian, Gary R. 33,875  
; REGISTRATION NUMBER: 8600-0139  
; REFERENCE/DOCKET NUMBER: 8600-0139  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 517 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
; US-08-190-802A-30

Query Match 100.0%; Score 172; DB 1; Length 517;  
Best Local Similarity 100.0%; Pred. No. 4,3e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Query 1 LHHCEAVLHLRFNNGMMVTCCKDRSLAVMD 31

Db 307 LIHCEAVLHLRFNNGMVTCSKDRSIAVWD 337

## RESULT 2

US-08-477-346-30  
Sequence 30, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
City: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-477-346-30

Query Match 100.0%; Score 172; DB 4; Length 517;

Best Local Similarity 100.0%; Pred. No. 4.3e-18;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIHCEAVLHLRFNNGMVTCSKDRSIAVWD 31  
Db 307 LIHCEAVLHLRFNNGMVTCSKDRSIAVWD 337

## RESULT 3

US-08-473-089-30  
Sequence 30, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
City: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-473-089-30

Query Match 100.0%; Score 172; DB 4; Length 517;

Best Local Similarity 100.0%; Pred. No. 4.3e-18;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIHCEAVLHLRFNNGMVTCSKDRSIAVWD 31  
Db 307 LIHCEAVLHLRFNNGMVTCSKDRSIAVWD 337

## RESULT 4

US-08-487-072A-30  
Sequence 30, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
City: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-487-072A-30

Query Match 100.0%; Score 172; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 4.3e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IHHCEAVLHLRFNNGMVTCSKDRSIAVMD 31  
Db 307 IHHCEAVLHLRFNNGMVTCSKDRSIAVMD 337

RESULT 5  
US-08-190-802A-85  
Sequence 85, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 85:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rIII, Fig. 13  
US-08-190-802A-85

Query Match 97.7%; Score 168; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 5.3e-19;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IHHCEAVLHLRFNNGMVTCSKDRSIAVMD 31

Db 1 IHHCEAVLHLRFNNGMVTCSKDRSIAVMD 30

RESULT 6  
US-08-477-346-85  
Sequence 85, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 85:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rIII, Fig. 13  
US-08-477-346-85

Query Match 97.7%; Score 168; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 5.3e-19;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IHHCEAVLHLRFNNGMVTCSKDRSIAVMD 31  
Db 1 IHHCEAVLHLRFNNGMVTCSKDRSIAVMD 30

RESULT 7  
US-08-473-089-85  
Sequence 85, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

```

1 STREET: 2000 Pennsylvania Avenue, NW
2 City: Washington
3 STATE: DC
4 COUNTRY: USA
5 ZIP: 20006-1812
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: PatentIn Release #1.0, Version #1.25
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/473,089
14 FILING DATE: 07-JUN-1995
15 CLASSIFICATION: 435
16 ATTORNEY/AGENT INFORMATION:
17 NAME: MURASHIGE, KATE H.
18 REGISTRATION NUMBER: 29,959
19 REFERENCE/DOCKET NUMBER: 2550-0025.22
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (202) 887-1500
22 TELEFAX: (202) 887-0763
23 INFORMATION FOR SEQ. ID NO.: 85:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 30 amino acids
26 TYPE: amino acid
27 TOPOLOGY: unknown
28 MOLECULE TYPE: peptide
29 HYPOTHEICAL: NO
30 ANTI-SENSE: NO
31 ORIGINAL SOURCE:
32 INDIVIDUAL ISOLATE: BETA TRCP rIII, Fig. 13
33
34 US-08-473-089-85
35
36 Query Match 97.7%; Score 168; DB 4; Length 30;
37 Matches Local Similarity 100.0%; Pred. NO. 5.3e-19;
38 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
39
40 QY 2 IHCEAVLHLRFNNGMMVTGSKRDSIAVWD 31
41 ||||||||||||||||||||||||||||
42 Db 1 IHCEAVLHLRFNNGMMVTGSKRDSIAVWD 30
43
44 RESULT 8
45 US-08-487-072a-85
46 Sequence 85, Application US/08487072A
47 Patent No. 6423684
48 GENERAL INFORMATION:
49 APPLICANT: Mochly-Rosen, Daria
50 APPLICANT: Ron, Dorit
51 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
52 TITLE OF INVENTION: Theroeof
53 NUMBER OF SEQUENCES: 265
54 CORRESPONDENCE ADDRESS:
55 ADDRESSEE: Morrison & Foerster
56 STREET: 2000 Pennsylvania Avenue, NW
57 City: Washington
58 STATE: DC
59 COUNTRY: USA
60 ZIP: 20006-1812
61 COMPUTER READABLE FORM:
62 MEDIUM TYPE: Floppy disk
63 COMPUTER: IBM PC compatible
64 OPERATING SYSTEM: PC-DOS/MS-DOS
65 SOFTWARE: PatentIn Release #1.0, Version #1.25
66 CURRENT APPLICATION DATA:
67 APPLICATION NUMBER: US/08/487,072A
68 FILING DATE: 07-JUN-1995
69 CLASSIFICATION: 514
70 ATTORNEY/AGENT INFORMATION:
71 NAME: MURASHIGE, KATE H.
72 REGISTRATION NUMBER: 29,959
73 REFERENCE/DOCKET NUMBER: 2550-0025.20
74 TELECOMMUNICATION INFORMATION:

```

[illegible]



COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/063,743  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Certone, Michael C  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0508 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LUNGMOT01  
CLONE: 059953  
US-09-063-743-1

Query Match 37.2% Score 64; DB 4; Length 375;  
Best Local Similarity 40.0%; Pred. No. 0.07;  
Matches 12; Conservative 7; Mismatches 9; Indels 2; Gaps 1;

QY 4 HCEAVLHLRFN--NGMMVTCSDRSIAVWD 31  
| | | | | : | | | | : | | | | |  
Db 108 HSGAVMELHYNTDGSMLFASDTKTVAVWD 137

RESULT 11  
US-09-590-540-1  
Sequence 1, Application US/09590540  
Patent No. 6410267  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
Tal, Preeti  
Guegler, Karl J.  
Corley, Neil C.  
Paterson, Chandra  
TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Genomics, Inc.  
STREET: 3160 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 3.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/590,540  
FILING DATE: 08-Jun-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/063,743  
FILING DATE: April 21, 1998

ATTORNEY/AGENT INFORMATION:  
NAME: David G. Streeter  
REGISTRATION NUMBER: 43,168  
REFERENCE/DOCKET NUMBER: PF-0508-1 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LUNGMOT01  
CLONE: 059953  
US-09-590-540-1

Query Match 37.2% Score 64; DB 4; Length 375;  
Best Local Similarity 40.0%; Pred. No. 0.07;  
Matches 12; Conservative 7; Mismatches 9; Indels 2; Gaps 1;

QY 4 HCEAVLHLRFN--NGMMVTCSDRSIAVWD 31  
| | | | | : | | | | : | | | | |  
Db 108 HSGAVMELHYNTDGSMLFASDTKTVAVWD 137

RESULT 12  
US-08-751-189-4  
Sequence 4, Application US/08751189  
Patent No. 5919656  
GENERAL INFORMATION:  
APPLICANT: Harrington, Lea A.  
Robinson, Murray O.  
TITLE OF INVENTION: No. 5919656e1 Genes Encoding Telomerase Protein  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen, Inc.  
STREET: 1840 De Havilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/751,189  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oleski, Nancy A.  
REGISTRATION NUMBER: 34,688  
REFERENCE/DOCKET NUMBER: A-433  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2629 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-751-189-4

Query Match 34.9% Score 60; DB 2; Length 2629;  
Best Local Similarity 39.3%; Pred. No. 3.1;  
Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMMVTCSDRSIAVWD 31

Db 2556 HIGSVTALHVLPGLLVTASKDRDKLWE 2583

## RESULT 13

US-09-060-836-4  
; Sequence 4, Application US/09060836  
; Patent No. 5981707  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, Lea A.  
; ATTORNEY/AGENT INFORMATION:  
; TITLE OF INVENTION: No. 5981707el Genes Encoding Telomerase Protein  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESSES: 12  
; ADDRESSEE: Amgen, Inc.  
; STREET: 1840 De Havilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/060,836  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/751,189  
; FILING DATE: 15-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oleski, Nancy A.  
; REGISTRATION NUMBER: 34,688  
; REFERENCE/DOCKET NUMBER: A-433  
; INFORMATION FOR SEO ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2629 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-060-836-4  
Query Match 34.9%; Score 60; DB 2; Length 2629;  
Best Local Similarity 39.3%; Pred. No. 3.1;  
Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;  
QY 4 HCEAVLHLRFNNGMVTCSKDRSIWVD 31  
Db 2556 HIGSVTALHVLPGLLVTASKDRDKLWE 2583  
RESULT 14  
US-09-184-445-4  
; Sequence 4, Application US/09184445  
; Patent No. 6174703  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, Lea A.  
; ATTORNEY/AGENT INFORMATION:  
; TITLE OF INVENTION: No. 6174703el Genes Encoding Telomerase Protein  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESSES: 12  
; ADDRESSEE: Amgen, Inc.  
; STREET: 1840 De Havilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/184,445  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/751,189  
; FILING DATE: 15-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oleski, Nancy A.  
; REGISTRATION NUMBER: 34,688  
; REFERENCE/DOCKET NUMBER: A-433  
; INFORMATION FOR SEO ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2629 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-184-445-4

Query Match 34.9%; Score 60; DB 4; Length 2629;  
Best Local Similarity 39.3%; Pred. No. 3.1;  
Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;  
QY 4 HCEAVLHLRFNNGMVTCSKDRSIWVD 31  
Db 2556 HIGSVTALHVLPGLLVTASKDRDKLWE 2583  
RESULT 15  
US-09-302-769-14  
; Sequence 14, Application US/09302769  
; Patent No. 6323317  
; GENERAL INFORMATION:  
; APPLICANT: HILTON, Douglas J  
; APPLICANT: ALEXANDER, Warren S  
; APPLICANT: VINEY, Elizabeth M  
; APPLICANT: WILSON, Tracey A  
; APPLICANT: RICHARDSON, Rachael T  
; APPLICANT: STARR, Robyn  
; APPLICANT: NICHOLSON, Sandra E  
; APPLICANT: METCALF, Donald  
; APPLICANT: NICOLA, Nicos A  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS  
; FILE REFERENCE: 109762  
; CURRENT APPLICATION NUMBER: US/09/302,769  
; CURRENT FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: 08/962,560  
; PRIOR FILING DATE: 1997-10-31  
; NUMBER OF SEO ID NOS: 50  
; SOFTWARE: Patentln Ver. 2.0  
; SEO ID NO 14  
; LENGTH: 421  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-302-769-14  
Query Match 34.0%; Score 58.5; DB 4; Length 421;  
Best Local Similarity 35.3%; Pred. No. 0.55;  
Matches 12; Conservative 9; Mismatches 10; Indels 3; Gaps 1;  
QY 1 LTHCEAVLHLRF---NNGMVTCSKDRSIWVD 31  
Db 166 LVDHTEWVDLTFAPDGSLLVSASRDKTLRWVD 199  
Search completed: April 10, 2003, 12:51:18  
Job time : 3.2095 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:49:06 ; Search time 3.17174 Seconds

(without alignments)  
597.532 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_343\_373  
Sequence: 172  
1 LIHCEAVLHLRFNNGMVTCSKDRSIAMVD 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	100.0	569	9	US-10-038-010-8
2	172	100.0	569	12	US-10-042-417-2
3	65	37.8	640	9	US-10-060-019-30
4	64	37.2	375	9	US-10-119-932-1
5	62	36.0	540	9	US-09-213-888-7
6	62	36.0	540	9	US-09-213-888-10
7	62	36.0	540	9	US-09-328-877A-7
8	62	36.0	540	9	US-09-328-877A-10
9	62	36.0	545	9	US-09-213-888-6
10	62	36.0	545	9	US-09-328-877A-6
11	62	36.0	553	9	US-09-213-888-5
12	62	36.0	553	9	US-09-328-877A-5
13	62	36.0	559	9	US-09-213-888-9
14	62	36.0	559	9	US-09-328-877A-9
15	62	36.0	589	9	US-09-213-888-8
16	62	36.0	589	9	US-09-328-877A-8
17	62	36.0	592	9	US-09-213-888-4
18	62	36.0	592	9	US-09-328-877A-4
19	62	36.0	626	9	US-09-213-888-21

20	62	36.0	626	9	US-09-328-877A-21	Sequence 21, Appl
21	62	36.0	627	9	US-09-213-888-3	Sequence 3, Appl1
22	62	36.0	627	9	US-09-328-877A-3	Sequence 27, Appl
23	62	36.0	666	9	US-09-213-888-27	Sequence 27, Appl
24	62	36.0	666	9	US-09-328-877A-27	Sequence 25, Appl
25	62	36.0	669	9	US-09-213-888-25	Sequence 25, Appl
26	62	36.0	669	9	US-09-328-877A-25	Sequence 2, Appl1
27	58.5	34.0	389	10	US-09-767-770A-2	Sequence 4, Appl1
28	58.5	34.0	421	10	US-09-767-770A-4	Sequence 14, Appl
29	58.5	34.0	421	10	US-09-808-805B-14	Sequence 314, App
30	58	33.7	678	10	US-09-801-368-314	Sequence 45365, A
31	57	33.1	130	10	US-09-864-761-45365	Sequence 2, Appl1
32	57	33.1	1194	10	US-09-876-667-2	Sequence 16, Appl
33	57	33.1	1205	10	US-09-876-667-16	Sequence 4, Appl
34	56.5	32.8	261	9	US-10-132-744A-4	Sequence 2, Appl1
35	56.5	32.8	484	9	US-10-132-744A-2	Sequence 6, Appl1
36	56.5	32.8	485	9	US-10-132-744A-6	Sequence 168, App
37	56	32.6	422	12	US-10-042-417-4	Sequence 449, App
38	56	32.6	498	10	US-09-729-674-168	Sequence 1262, Ap
39	55	32.0	521	10	US-09-764-853-449	Sequence 1085, Ap
40	55	32.0	732	9	US-09-832-292-12	Sequence 46317, A
41	54	31.4	90	10	US-09-867-550-1262	Sequence 2054, Ap
42	53.5	31.1	392	10	US-09-925-300-1085	Sequence 221, App
43	53	30.8	50	10	US-09-864-761-46317	
44	53	30.8	103	10	US-09-867-550-2054	
45	52.5	30.5	476	9	US-10-001-835-221	

## ALIGNMENTS

RESULT 1  
US-10-038-010-8  
; Sequence 8, Application US/10038010  
; Publication No. US20030040089A1  
; GENERAL INFORMATION:  
; APPLICANT: HYBRIGENICS  
; APPLICANT: Pierre, Legrain  
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells  
; FILE REFERENCE: B4767A  
; CURRENT APPLICATION NUMBER: US/10/038,010  
; CURRENT FILING DATE: 2002-07-23  
; PRIOR APPLICATION NUMBER: US 60/259,377  
; PRIOR FILING DATE: 2001-01-02  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 8  
; LENGTH: 569  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: beta-Trcp1  
; LOCATION: (1)..(569)  
; OTHER INFORMATION:  
US-10-038-010-8

Query Match 100.0%; Score 172; DB 9; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.1e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHCEAVLHLRFNNGMVTCSKDRSIAMVD 31  
DB 343 LIHCEAVLHLRFNNGMVTCSKDRSIAMVD 373

RESULT 2  
US-10-042-417-2  
; Sequence 2, Application US/10042417  
; Patent No. US20020123082A1  
; GENERAL INFORMATION:  
; APPLICANT: Pagano, M.  
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS

FILE REFERENCE: 5914-090-999  
CURRENT APPLICATION NUMBER: US/10/042,417  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-042-417-2

Query Match 100.0%; Score 172; DB 12; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1,1e-17;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LHHCEAVLHLRFNNGMVTGSKDRSIAVWD 31  
DB 343 LHHCEAVLHLRFNNGMVTGSKDRSIAVWD 373

RESULT 3  
US-10-060-019-30  
Sequence 30, Application US/10060019  
Publication No. US20030003564A1  
GENERAL INFORMATION:  
APPLICANT: Tyers, Mike  
APPLICANT: Williams, Andrew  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN  
FILE REFERENCE: 11/57,100SU01  
CURRENT APPLICATION NUMBER: US/10/060,019  
PRIOR FILING DATE: 2002-01-29  
PRIOR APPLICATION NUMBER: US/09/177,165  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: 60/092,443  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/063,254  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 30  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-060-019-30

Query Match 37.8%; Score 65; DB 9; Length 640;  
Best Local Similarity 42.9%; Pred. No. 0.12;  
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMVTGSKDRSIAVWD 31  
DB 301 HMDGVLTLOFNRLFTGSDYSTICIMD 328

RESULT 4  
US-10-119-932-1  
Sequence 1, Application US/10119932  
Patent No. US20020160488A1  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Genomics, Inc.  
STREET: 3160 Porter Drive  
CITY: Palo Alto

STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 3.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/119,932  
FILING DATE: 09-Apr-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/590,540  
FILING DATE: 08-Jun-2000  
APPLICATION NUMBER: 09/063,743  
FILING DATE: April 21, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: David G. Streeter

REGISTRATION NUMBER: 43,168  
REFERENCE/DOCKET NUMBER: PF-0508-1 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LUNGN0701  
CLONE: 059953

SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-119-932-1

Query Match 37.2%; Score 64; DB 9; Length 375;  
Best Local Similarity 40.0%; Pred. No. 0.094;  
Matches 12; Conservative 7; Mismatches 9; Indels 2; Gaps 1;

QY 4 HCEAVLHLRFN--NGMVTGSKDRSIAVWD 31  
DB 108 HSGAVMELHYNTDGSMLFSATDXTVAVWD 137

RESULT 5  
US-09-213-888-7  
Sequence 7, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-7

Query Match 36.0%; Score 62; DB 9; Length 540;  
Best Local Similarity 39.3%; Pred. No. 0.28;  
Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMVTGSKDRSIAVWD 31

Db 456 HOSAVTCLQFNKNFVITSSDDGTVKLM 483

RESULT 6

US-09-213-888-10  
; Sequence 10, Application US/09213888A  
; Patent No. US20020164683A1

GENERAL INFORMATION:

APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe

APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company

TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/213,888A

CURRENT FILING DATE: 1998-12-17

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 10

LENGTH: 540

TYPE: PRT

ORGANISM: Homo sapiens

US-09-213-888-10

Query Match 36.0%; Score 62; DB 9; Length 540;

Best Local Similarity 39.3%; Pred. No. 0.28;

Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 4 HCEAVLHLRFNNGMVTCSKDRSIAVMD 31

Db 456 HOSAVTCLQFNKNFVITSSDDGTVKLM 483

RESULT 7

US-09-328-877A-7  
; Sequence 7, Application US/09328877A  
; Patent No. US20020177187A1

GENERAL INFORMATION:

APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe

APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company

TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/328,877A

CURRENT FILING DATE: 1999-06-09

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7

LENGTH: 540

TYPE: PRT

ORGANISM: Homo sapiens

US-09-328-877A-7

Query Match 36.0%; Score 62; DB 9; Length 540;

Best Local Similarity 39.3%; Pred. No. 0.28;

Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 4 HCEAVLHLRFNNGMVTCSKDRSIAVMD 31

Db 456 HOSAVTCLQFNKNFVITSSDDGTVKLM 483

RESULT 8

US-09-328-877A-10  
; Sequence 10, Application US/09328877A  
; Patent No. US20020177187A1

GENERAL INFORMATION:

APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe

APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company

TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/328,877A

CURRENT FILING DATE: 1999-06-09

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 10

LENGTH: 540

TYPE: PRT

ORGANISM: Homo sapiens

US-09-328-877A-10

Query Match 36.0%; Score 62; DB 9; Length 540;

Best Local Similarity 39.3%; Pred. No. 0.28;

Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 4 HCEAVLHLRFNNGMVTCSKDRSIAVMD 31

Db 456 HOSAVTCLQFNKNFVITSSDDGTVKLM 483

RESULT 9

US-09-213-888-6  
; Sequence 6, Application US/09213888A  
; Patent No. US20020164683A1

GENERAL INFORMATION:

APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe

APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company

TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/213,888A

CURRENT FILING DATE: 1998-12-17

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 6

LENGTH: 545

TYPE: PRT

ORGANISM: Homo sapiens

US-09-213-888-6

Query Match 36.0%; Score 62; DB 9; Length 545;

Best Local Similarity 39.3%; Pred. No. 0.28;

Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 4 HCEAVLHLRFNNGMVTCSKDRSIAVMD 31

Db 461 HOSAVTCLQFNKNFVITSSDDGTVKLM 488

RESULT 10

US-09-328-877A-6  
; Sequence 6, Application US/09328877A  
; Patent No. US20020177187A1

GENERAL INFORMATION:

APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe

APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company

TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/328,877A

CURRENT FILING DATE: 1999-06-09

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 6

LENGTH: 545

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-6

Query Match  
Best Local Similarity 36.0%; Score 62; DB 9; Length 545;  
Matches 11: Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMAYTCSKDRSIAVWD 31  
DB 461 HOSAVTCLQFNKNFVITSSDDGTVKLMD 488

RESULT 11  
US-09-213-888-5  
Sequence 5, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-5

Query Match  
Best Local Similarity 39.3%; Pred. No. 0.29;  
Matches 11: Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMAYTCSKDRSIAVWD 31  
DB 469 HOSAVTCLQFNKNFVITSSDDGTVKLMD 496

RESULT 12  
US-09-328-877A-5  
Sequence 5, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-5

Query Match  
Best Local Similarity 36.0%; Score 62; DB 9; Length 553;  
Matches 11: Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMAYTCSKDRSIAVWD 31  
DB 469 HOSAVTCLQFNKNFVITSSDDGTVKLMD 496

RESULT 13  
US-09-213-888-9

Sequence 9, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 559  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-9

Query Match  
Best Local Similarity 36.0%; Score 62; DB 9; Length 559;  
Matches 11: Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMAYTCSKDRSIAVWD 31  
DB 475 HOSAVTCLQFNKNFVITSSDDGTVKLMD 502

RESULT 14  
US-09-328-877A-9  
Sequence 9, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 559  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-9

Query Match  
Best Local Similarity 39.3%; Pred. No. 0.29;  
Matches 11: Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNGMAYTCSKDRSIAVWD 31  
DB 475 HOSAVTCLQFNKNFVITSSDDGTVKLMD 502

RESULT 15  
US-09-213-888-8  
Sequence 8, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
US-09-213-888-8

```

; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 8
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-8

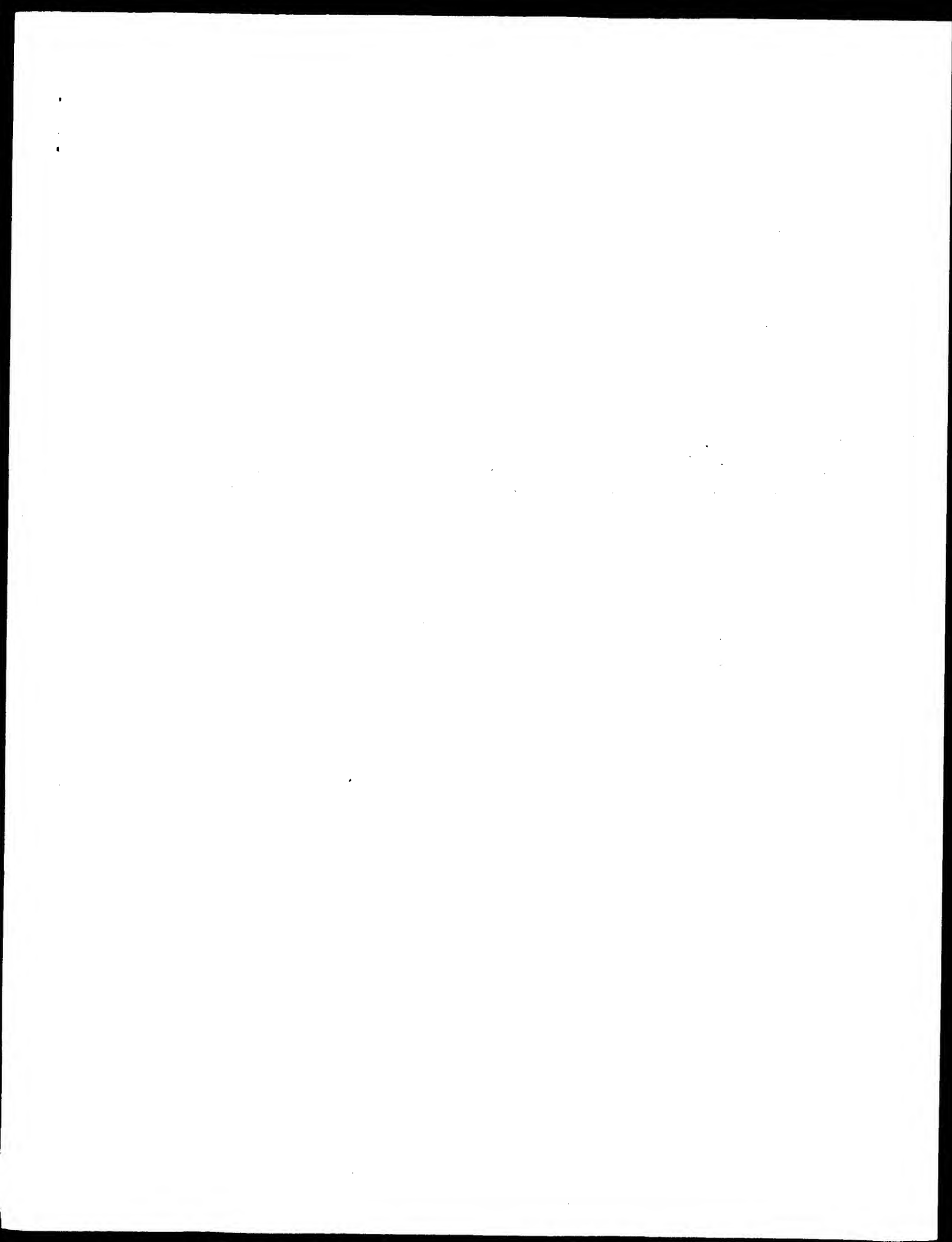
```

```

Query Match          36.0%; Score 62; DB 9; Length 589;
Best Local Similarity 39.3%; Pred. NO. 0.31;
Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
QY      4 HCEAVLHLRFNNGMMVTCCKDRSIKAWD 31
        | | | | | | | | | | | | | | | |
DB      505 HQSAVTCTIQFNKNEFYITSSDDGTAKLWD 532

```

Search completed: April 10, 2003, 13:16:45  
Job time : 4.17174 secs





GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 3.58709 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_343\_373

Sequence: 172  
1 LIHCEAVLHLRFNNGMMVTCKDRSIAVWD 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	172	100.0	518	2 B48088	beta-transducin re
2	160	93.0	701	2 T16607	hypothetical prote
3	71	41.3	714	2 S56893	hypothetical prote
4	67	39.0	1893	2 T22661	hypothetical prote
5	66	38.4	331	2 T20593	hypothetical prote
6	65	37.8	640	2 S49932	MET30 protein - ye
7	64	37.2	659	2 S38108	hypothetical prote
8	62.5	36.3	810	2 T48835	lethal(2)denticlel
9	62	36.0	506	2 T50211	WD-repeat protein
10	62	36.0	507	2 T38653	trp-asp repeat pro
11	61	35.5	1013	2 S65195	probable membrane
12	60.5	35.2	506	2 S50937	hypothetical prote
13	60	34.9	303	2 S45461	hypothetical prote
14	60	34.9	2629	2 T30987	telomerase-associ
15	59	34.3	315	2 S46046	probable GRP-bindi
16	59	34.3	351	2 C88082	protein T05A8.5 [l
17	59	34.3	415	2 T20335	hypothetical prote
18	59	34.3	431	2 T38471	probable chromatin
19	58.5	34.0	421	2 T14773	hypothetical prote
20	58	33.7	1693	2 S76086	beta transducin-11
21	57.5	33.4	490	2 T37884	transcription fact
22	57.5	33.4	576	2 S29367	PMF1 protein - yea
23	57	33.1	325	2 T06784	GRP-binding protei
24	57	33.1	431	2 A45142	cleavage stimulat
25	57	33.1	561	2 T19604	hypothetical prote
26	57	33.1	651	2 T50289	WD repeat protein
27	57	33.1	1194	2 T03818	apoptotic proteina
28	57	33.1	1227	2 AE810	WD-40 repeat prote
29	56.5	32.8	793	2 T27133	hypothetical prote

30	56	32.6	64	2 T04116	GRP-binding protei
31	56	32.6	326	2 T16987	GRP-binding protei
32	56	32.6	326	2 T02340	GRP-binding regula
33	56	32.6	384	2 S46110	hypothetical prote
34	56	32.6	422	2 A56640	CDC4 repeat unit-c
35	56	32.6	430	2 T41054	probable chromosom
36	56	32.6	430	2 T21482	hypothetical prote
37	56	32.6	454	2 T01337	hypothetical prote
38	56	32.6	465	2 S51295	hypothetical prote
39	56	32.6	465	2 T00982	En/Spm-like transp
40	56	32.6	515	2 S19487	hypothetical prote
41	56	32.6	629	2 T48799	Rip9p related prot
42	56	32.6	2629	2 T32735	telomerase-associ
43	55.5	32.3	243	2 T05307	hypothetical prote
44	55	32.0	82	2 T31574	hypothetical prote
45	55	32.0	267	2 S62507	hypothetical trp-a

## ALIGNMENTS

## RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-Trcp  
C:Species: Xenopus laevis (African clawed frog)

C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000

C:Accession: B48088

R:Spewak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: A48088; MUID:9330269; PMID:8393141

A:Accession: B48088

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SPC>

A:Cross-References: GB:W98268; NID:g295542; PIDN:AAA02810.1; PID:g295543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <WD1>

## Query Match

Score 172; DB 2; Length 518;

Best Local Similarity 100.0%; Pred. No. 3.2e-17;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHCEAVLHLRFNNGMMVTCKDRSIAVWD 31

DB 308 LIHCEAVLHLRFNNGMMVTCKDRSIAVWD 338

## RESULT 2

T16607 hypothetical protein K10B2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16607

R:Miller, N. submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid K10B2.

A:Reference number: Z18545

A:Accession: T16607

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-701 <MII>

A:Cross-References: EMBL:U28730; NID:g860694; PID:g860695; PIDN:AAA68258.1; CESP:K10B

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K10B2.1

A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

## Query Match

Score 160; DB 2; Length 701;

Best Local Similarity 93.5%; Pred. No. 2.4e-15;

Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LIHCEAVLHLRFNNNGMYTCSKDRSTAVMD 31  
 |||||  
 Db 334 LIHCEAVLHLRFNNNGMYTCSKDRSTAVMD 364

## RESULT 3

S56893  
 hypothetical protein YKR036c homolog YJL112w - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein J0802

C:Species: Saccharomyces cerevisiae

C>Date: 05-May-1995 #sequence\_revision 08-Sep-1995 #text\_change 19-Apr-2002

C:Accession: S56893; S56890

R:Ciepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.

submitted to the Protein Sequence Database, September 1995

A:Reference number: S56891

A:Accession: S56893

A:Molecule type: DNA

A:Residues: 1-714 <CZ1>

A:Cross-references: EMBL:Z49387; NID:g1008298; PIDN:CAA89407.1; PID:g1008299; MIPS:YJL112w

R:Rasmussen, S.W.

submitted to the Protein Sequence Database, September 1995

A:Reference number: S56876

A:Accession: S56890

A:Molecule type: DNA

A:Residues: 597-714 <RAS>

A:Cross-references: EMBL:Z49387; MIPS:YJL112w

C:Genetics:

A:Gene: SGD:MDV1

A:Cross-references: SGD:S0003648

A:Map position: 10L

C:Superfamily: yeast hypothetical protein YKR036c; WD repeat homology

F:437-470/Domain: WD repeat homology <WD1>

F:498-529/Domain: WD repeat homology <WD2>

F:602-633/Domain: WD repeat homology <WD3>

Query Match

Best Local Similarity 41.3%; Score 71; DB 2; Length 714;

Matches 11; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNNGMYTCSKDRSTAVMD 31

Db 605 HFDATITSLKFDASCLVGTSDRTVRIMD 632

hypothetical protein T01G1.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T22661; T24305

R:White, S.; Mortimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19594

A:Accession: T22661

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1893 <WIL>

A:Cross-references: EMBL:Z82271; PIDN:CAB05214.1; GSPDB:GN00022; CESP:T01G1.1

A:Experimental source: clone F54E12

R:Kershaw, J.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19872

A:Accession: T24305

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1893 <WIL>

A:Cross-references: EMBL:Z22811; PIDN:CAB07273.1; GSPDB:GN00022; CESP:T01G1.1

A:Experimental source: clone T01G1

C:Genetics:

A:Gene: CESP:T01G1.1

A:Map position: 4

A:Introns: 65/3; 101/3; 194/3; 269/3; 299/2; 373/3; 481/3; 576/3; 615/2; 704/3; 839/1; 9

Query Match

Best Local Similarity 39.0%; Score 67; DB 2; Length 1893;

Matches 13; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 4 HCEAVLHLRFNNNGMYTCSKDRSTAVMD 31

Db 1571 HARGVSLVDVNEKLMVTGSKDRTRAKIMD 1598

hypothetical protein F08G12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 26-May-2000

C:Accession: T20593

R:Cottage, A.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19297

A:Accession: T20593

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-331 <WIL>

A:Cross-references: EMBL:Z66561; PIDN:CAA91460.1; GSPDB:GN00028; CESP:F08G12.2

A:Experimental source: clone F08G12

C:Genetics:

A:Gene: CESP:F08G12.2

A:Map position: X

A:Introns: 6/1; 28/3; 61/3; 314/1

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match

Best Local Similarity 38.4%; Score 66; DB 2; Length 331;

Matches 12; Conservative 8; Mismatches 8; Indels 2; Gaps 1;

QY 4 HCEAVLHLRF--NNGMYTCSKDRSTAVMD 31

Db 80 HSGAVMDLKFTRDSSSLVSCGTRKSVRYMD 109

MET30 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein Y1905.02; protein Y1L046w

C:Species: Saccharomyces cerevisiae

C>Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 26-May-2000

C:Accession: S49932; S43750

R:Odell, C.; Bowman, S.

submitted to the EMBL Data Library, December 1994

A:Reference number: S49931

A:Accession: S49932

A:Molecule type: DNA

A:Residues: 1-640 <ODE>

A:Cross-references: GB:Z47047; EMBL:Z46861; NID:G603997; PID:g763300; MIPS:Y1L046w

R:Thomas, D.; Cherest, H.; Barbey, R.; Surdin-Kerjan, Y.

submitted to the EMBL Data Library, December 1993

A:Reference number: S43750

A:Accession: S43750

A:Molecule type: DNA

A:Residues: 1-60, 'I', 62-640 <THO>

A:Cross-references: EMBL:126505; NID:g432493; PID:g432494

C:Genetics:

A:Gene: SGD:MET30

A:Cross-references: SGD:S0001308; MIPS:Y1L046w

A:Map position: 9L

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

F:298-329/Domain: WD repeat homology <WD1>

F:338-369/Domain: WD repeat homology <WD2>

F:374-409/Domain: WD repeat homology <WD3>

F:417-450/Domain: WD repeat homology <WD4>

Query Match

Best Local Similarity 37.8%; Score 65; DB 2; Length 640;

Matches 12; Conservative 42.9%; Pred. No. 0.17;









GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:51:26 ; Search time 2.22777 Seconds

(Without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_343\_373  
Perfect score: 172  
Sequence: 1 LITHCEAVILHLEFNGMAYTCSKRSIAVMD 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	100.0	518	1 TRCB_XENLA	Q91854 xenopus lae
2	172	100.0	605	1 FWIA_HUMAN	Q9Y297 homo sapien
3	160	93.0	665	1 LI23_CABEL	Q09990 caenorhabd
4	152	88.4	542	1 FWIB_HUMAN	Q9UKB1 homo sapien
5	71	41.3	714	1 YJL2_YEAST	P47070 saccharomyc
6	65	37.8	640	1 MTG3_YEAST	P39014 saccharomyc
7	64	37.2	659	1 YK16_YEAST	P36130 saccharomyc
8	62	36.0	506	1 POPE_SCHPO	Q09855 schizosacch
9	60	34.9	303	1 YNA6_YEAST	P41318 saccharomyc
10	59	34.3	315	1 YB25_YEAST	P38123 saccharomyc
11	59	34.3	415	1 YLN2_CABEL	Q18964 caenorhabd
12	59	34.3	431	1 YDPT_SCHPO	Q14021 schizosacch
13	59	34.3	1249	1 APAF_RAT	Q99P51 rattus norv
14	58.5	34.0	421	1 WSBI_HUMAN	Q9Y617 homo sapien
15	58.5	34.0	421	1 WSBI_MOUSE	Q54927 mus musculu
16	58	33.7	678	1 SCOB_EMENT	Q00659 emeritella
17	58	33.4	1693	1 Y163_SYNY3	Q55563 synecocyst
18	57.5	33.4	490	1 YEOJ_SCHPO	Q10990 schizosacch
19	57.5	33.4	576	1 PMP1_YEAST	P21304 saccharomyc
20	57	33.1	325	1 GBIP_SOYBN	Q39856 glycine max
21	57	33.1	431	1 CST1_HUMAN	Q05048 homo sapien
22	57	33.1	651	1 YIX1_SCHPO	Q9P713 schizosacch
23	57	33.1	1248	1 APAF_HUMAN	Q14727 homo sapien
24	56	33.1	1249	1 APAF_MOUSE	Q88879 mus musculu
25	56	32.6	326	1 GBIP_NICPL	P93340 nicotiana p
26	56	32.6	326	1 GBIP_TOBAC	P49026 nicotiana t
27	56	32.6	384	1 AR41_YEAST	P38328 saccharomyc
28	56	32.6	422	1 FBW2_HUMAN	Q9UKC8 homo sapien
29	56	32.6	422	1 FBW2_MOUSE	Q60584 mus musculu
30	56	32.6	465	1 YN57_YEAST	P42841 saccharomyc
31	56	32.6	515	1 YCWM2_YEAST	P25382 saccharomyc
32	56	32.6	664	1 EMAP_LYRYA	Q9Y1C1 lytechinus
33	55	32.0	922	1 YB1C_SCHPO	P67177 schizosacch

34	54.5	31.7	446	1 GRWD_HUMAN	Q9B67 homo sapien
35	54	31.4	318	1 PEX7_MOUSE	P97865 mus musculu
36	54	31.4	325	1 GBIP_MENSA	Q24076 medicago sa
37	54	31.4	327	1 GBIP_ARATH	Q24456 arabidopsis
38	54	31.4	327	1 GBIP_BRANA	Q29336 brassica na
39	54	31.4	339	1 CIA1_HUMAN	Q76071 homo sapien
40	54	31.4	455	1 PR17_YEAST	P40968 saccharomyc
41	54	31.4	872	1 VP2_ROTTC	P26191 porcine rot
42	53.5	31.1	349	1 YDFU_ECOLI	P76162 escherichia
43	53.5	31.1	519	1 TBL3_HUMAN	Q12768 homo sapien
44	53	30.8	318	1 GBIP_DROME	O18640 drosophila
45	53	30.8	323	1 PEX7_HUMAN	Q00628 homo sapien

## ALIGNMENTS

```

RESULT 1
ID TRCB_XENLA STANDARD: PRT: 518 AA.
AC Q91854: P70037: P70038:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-TrCP (Beta-transducin repeat-containing protein).
GN FBXW1 OR BTRCP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93330289; PubMed=8393141;
RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
RT anaphase are rescued by Xenopus CDNs encoding N-ras or a protein
RT with beta-transducin repeats."
RL Mol. Cell. Biol. 13:4953-4966(1993).
RN [2]
RP SEQUENCE OF 302-518 FROM N.A.
RX MEDLINE=97109804; PubMed=8952061;
RA Hudson J.W., Alatron V.B., Elinson R.P.;
RT "Identification of new localized RNAs in the Xenopus oocyte by
RT differential display PCR."
RL Dev. Genet. 19:190-198(1996).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -!- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIBLASTULA TRANSITION. DO
CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
CC TADPOLE EMBRIO.
CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M98268; AAA02810.1; -
CC EMBL: U63921; AAB49671.1; -
CC EMBL: U63922; AAB49672.1; -
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00400; WD40; 7.

```

DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ubl conjugation pathway; Repeat; WD repeat.  
 KW DOMAIN 119 157 F-BOX.  
 FT REPEAT 230 258 WD 1.  
 FT REPEAT 270 298 WD 2.  
 FT REPEAT 310 338 WD 3.  
 FT REPEAT 353 381 WD 4.  
 FT REPEAT 393 421 WD 5.  
 FT REPEAT 433 461 WD 6.  
 FT REPEAT 482 510 WD 7.  
 FT CONFLICT 302 304 GEM -> EFR (IN REF. 2).  
 FT CONFLICT 516 518 GLA -> AAH (IN REF. 2).  
 SQ SEQUENCE 518 AA; 59507 MW; 2A52EC19028127F3 CRC64;

Query Match 100.0%; Score 172; DB 1; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 LIHCEAVLHLRFNNGMVTCSKDRSTAVWD 31  
 Db 308 LIHCEAVLHLRFNNGMVTCSKDRSTAVWD 338

RESULT 2  
 ID FWLA\_HUMAN STANDARD; PRT; 605 AA.  
 AC G9Y297; G9Y213;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-Trcp)  
 DE (3KR51kappa) (PikappaBalpha-E3 receptor subunit).  
 GN FBXW1A OR FBW1A OR BTBRCP OR BTRC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=9075339; PubMed=9859996;  
 RA Varon A., Hatzidai A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
 RT "Identification of the receptor component of the IkappaBalpha-  
 RT ubiquitin ligase";  
 RL Nature 396:590-594(1998).  
 RN [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Lymphoid;  
 RX MEDLINE=98325370; PubMed=9660940;  
 RA Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,  
 RA Thomas D., Strebel K., Benarous R.;  
 RT "A novel human WD protein, h-beta Trcp, that interacts with HIV-1 Vpu  
 RT connects C4 to the ER degradation pathway through an F-box motif";  
 RL Mol. Cell 1:565-574(1998).  
 RN [3]  
 RN SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=20003060; PubMed=10531935;  
 RA Cenciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [4]  
 RN CHARACTERIZATION.  
 RX MEDLINE=99145464; PubMed=9990852;  
 RA Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,

RA Harper J.W.;  
 RT "The SCF(beta-Trcp)-ubiquitin ligase complex associates specifically  
 RT with phosphorylated destruction motifs in I-kappa-B-alpha and  
 RT beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro";  
 RL Genes Dev. 13:270-283(1999).  
 CC -!- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA  
 CC (PIKAPPABALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR  
 CC UBIQUITINATION AND DEGRADATION.  
 CC -!- SUBUNIT: PART OF A SCF (SKP1-CULIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

DR EMBL; AF101784; AAD08702.1; -  
 DR EMBL; Y14153; CA74572.1; -  
 DR EMBL; AF129530; AAP04464.1; -  
 DR Genew; HGNC:1144; BTRC.  
 DR MIM; 603482; -  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF004400; WD40; 7.  
 DR Pfam; PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 FT DOMAIN 190 228 F-BOX.  
 FT REPEAT 301 338 WD 1.  
 FT REPEAT 341 378 WD 2.  
 FT REPEAT 381 418 WD 3.  
 FT REPEAT 424 461 WD 4.  
 FT REPEAT 464 503 WD 5.  
 FT REPEAT 505 541 WD 6.  
 FT REPEAT 553 590 WD 7.  
 FT VANSPLC 17 52 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 605 AA; 68866 MW; 4C67F3B7E400FD37 CRC64;  
 Query Match 100.0%; Score 172; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 LIHCEAVLHLRFNNGMVTCSKDRSTAVWD 31  
 Db 379 LIHCEAVLHLRFNNGMVTCSKDRSTAVWD 409  
 RESULT 3  
 ID LI23\_CAEEL STANDARD; PRT; 665 AA.  
 AC Q09990; Q9GN6;  
 DT 15-JUN-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein lin-23.  
 GN LIN-23 OR K10B2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.



OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.: FUNCTION, DEVELOPMENTAL STAGE, AND MUTAGENESIS.  
 RX MEDLINE=20515608; PubMed=11060233;  
 RA Kipreos E.T., Gohel S.P., Hedgecock E.M.;  
 RT "The *Caenorhabditis elegans* F-box/WD-repeat protein Lin-23 functions  
 RL to limit cell division during development.";  
 RN Development 127:5071-5083(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Miller N.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Functions cell autonomously to negatively regulate cell  
 CC cycle progression. Required to restrain cell proliferation in  
 CC response to developmental cues. Probably recognizes and binds to  
 CC some proteins and promotes their ubiquitination and degradation  
 CC (By similarity).  
 CC -1- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex  
 CC (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- DEVELOPMENTAL STAGE: Highest levels in embryos and adults, lowest  
 CC levels in larvae. Maternal expression results in high zygotic  
 CC levels.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO X.LAEVIS FBXW1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF275253; AAC28037.1; -;  
 DR EMBL: U28730; AAA68258.2; -;  
 DR WormRep: K10B2.1; CE28600.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF004400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR PROSITE: PS00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS00982; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
 KW Developmental protein; Cell cycle; Cell division;  
 KW Ub1 conjugation pathway; Repeat; WD repeat.  
 FT DOMAIN 81 127  
 FT REPEAT 220 257  
 FT REPEAT 260 299  
 FT REPEAT 301 337  
 FT REPEAT 343 380  
 FT REPEAT 383 420  
 FT REPEAT 423 460  
 FT REPEAT 472 509  
 FT DOMAIN 525 582  
 FT MOTAEN 441 441  
 FT SEQUENCE 665 AA; 75916 MW; BF3E9AF51F12ECCC CRC64;  
 G->R: IN LIN-32(RH293).  
 Query Match 93.0%; Score 160; DB 1; Length 665;  
 Best Local Similarity 93.5%; Pred. No. 1,8e-16;  
 Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LIIHCEAVLHLRFNNGMMVYTSCKSRISIAWMD 31  
 DB 298 LIIHCEAVLHLRFNNGMMVYTSCKSRISIAWMD 328  
 RESULT 4  
 FFWB\_HUMAN STANDARD: PRT: 542 AA.  
 AC Q9UKB1: Q9Y4C6; Q9P2S8; Q9P2S9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TRCP2).  
 GN FBXW1B OR FBW1B OR BTRCP2 OR KIAA0696.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Cenciarelli C., Chiari D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Petal lung;  
 RX MEDLINE=20160458; PubMed=10694485;  
 RA Kohle J., Sagara N., Kitikoshi H., Takagi A., Miwa T., Hirai M.,  
 RA Katoh M.;  
 RT "Molecular cloning and genomic structure of the betaTRCP2 gene on  
 RT chromosome 5q35.1.";  
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B AND C (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF176022; AAF04528.1; -;  
 DR EMBL: AB033279; BAA92329.1; -;  
 DR EMBL: AB033280; BAA92330.1; -;  
 DR EMBL: AB033281; BAA92331.1; -;  
 DR EMBL: AB014596; BAA91671.1; ALT\_INIT.  
 DR Genew: HGNC:13607; FBXW1B.  
 DR MIM: 605651; -;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF004400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.





RA Shpakovski G.V., Usersty D., Barrell B.G., Nurse P.;  
 RT "the genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: CONTAINS 1 P-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AB061694; BAB55543.1; -  
 CC EMBL: AL136538; CAB6464.1; -  
 CC EMBL: Z66525; CAA91423.1; -  
 CC InterPro: IPR001810; F-box.  
 CC InterPro: IPR001680; WD40.  
 CC Pfam: PF00400; WD40; 7.  
 CC Pfam: PF00646; F-box; 1.  
 CC PRINTS: PR00320; GPROTEINBRPT.  
 CC ProDom: PD000018; WD40; 3.  
 CC SMART: SM00256; FBOX; 1.  
 CC SMART: SM00320; WD40; 8.  
 CC PROSITE: PS50181; FBOX; 1.  
 CC PROSITE: PS00678; WD\_REPEATS\_1; 3.  
 CC PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 CC PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT DOMAIN 70 116 F-BOX.  
 FT REPEAT 219 256 WD 1.  
 FT REPEAT 259 298 WD 2.  
 FT REPEAT 301 338 WD 3.  
 FT REPEAT 345 386 WD 4.  
 FT REPEAT 388 426 WD 5.  
 FT REPEAT 427 464 WD 6.  
 FT REPEAT 468 505 WD 7.  
 SQ SEQUENCE 506 AA; 58257 MW; CEF34D4EFFBC2E10 CRC64;  
 Query Match 36.0%; Score 62; DB 1; Length 506;  
 Best Local Similarity 38.5%; Pred. No. 0.086;  
 Matches 10; Conservative 12; Mismatches 4; Indels 0; Gaps 0;  
 QY 6 EAVLHRENGMNVTCSDRSIAVWD 31  
 DB 222 DSVYCVQYDEITVSGSKDRIVSWD 247  
 ID YNA6\_YEAST STANDARD; PRT; 303 AA.  
 AC P41318;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 34.0 kDa Trp-Asp repeats containing protein in SISI-MRLP2  
 DE intergenic region.  
 DE YNL006W OR N2005.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / FY1679;  
 RX MEDLINE=95076713; PubMed=7985421;  
 RA Verhasselt P., Aert R., Voet M., Volckaert G.;  
 RT "Nucleotide sequence analysis of an 8887 bp region of the left arm of  
 RL yeast chromosome XIV, encompassing the centromere sequence.";  
 RN [2]  
 RP SEQUENCE FROM N.A.

RA Dolignon F., Crouzet M.;  
 RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X77114; CAA54380.1; -  
 CC EMBL: Z71282; CAA95865.1; -  
 CC PIR: S45123; S45123.  
 CC PIR: S45461; S45461.  
 CC SGD: S0004951; YNL006W.  
 CC InterPro: IPR001680; WD40.  
 CC Pfam: PF00400; WD40; 7.  
 CC PRINTS: PR00320; GPROTEINBRPT.  
 CC ProDom: PD000018; WD40; 2.  
 CC SMART: SM00320; WD40; 6.  
 CC PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 CC PROSITE: PS50082; WD\_REPEATS\_2; 5.  
 CC PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 73 103 WD 1.  
 FT REPEAT 114 144 WD 2.  
 FT REPEAT 157 187 WD 3.  
 FT REPEAT 205 235 WD 4.  
 FT REPEAT 248 278 WD 5.  
 SQ SEQUENCE 303 AA; 34034 MW; 536AA1DB8E8151F0 CRC64;  
 Query Match 34.9%; Score 60; DB 1; Length 303;  
 Best Local Similarity 38.5%; Pred. No. 0.099;  
 Matches 10; Conservative 7; Mismatches 7; Indels 2; Gaps 1;  
 QY 6 EAVLHRENGMNVTCSDRSIAVWD 31  
 DB 121 EVVLIHP--NOGELISCDRDGIRIWD 144  
 ID YB25\_YEAST STANDARD; PRT; 315 AA.  
 AC P38123;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 34.8 kDa Trp-Asp repeats containing protein in SMY2-RPS6B  
 DE intergenic region.  
 DE YBR175W OR YBR1237.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RA Entian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,  
 RA Niegemann E., Schenk-Groeninger R., Servis J., Wehner E.,  
 RA Wolter R., Brendel M., Bauer J., Braun H., Dern K., Duesterhus S.,  
 RA Gruenbein R., Hedger D., Klesau P., Korol S., Krems B., Proft M.,  
 RA Siegers K., Baur A., Boles E., Miosga T.,  
 RA Schaff-Gerstenschlaeger I., Zimmermann F.K.;  
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z36043; CAA85136.1; -  
 DR PIR: S46046; S46046.  
 DR SGD: S0000379; YBR175W.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 5.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 1.  
 DR SMART: SM00320; WD40; 4.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 3.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein: Repeat; WD repeat.  
 FT REPEAT 53 93 WD 1.  
 FT REPEAT 94 133 WD 2.  
 FT REPEAT 136 178 WD 3.  
 FT REPEAT 187 228 WD 4.  
 FT REPEAT 238 278 WD 5.  
 FT REPEAT 285 314 WD 6.  
 SQ SEQUENCE 315 AA; 34753 MW; E8E8BC483D1CD776 CRC64;

Query Match 34.3%; Score 59; DB 1; Length 315;  
 Best Local Similarity 40.6%; Pred. No. 0.15;  
 Matches 13; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

OY 2 IHHCVALHLRFN-NGMAYTCSKRSIAVMD 31  
 Db 93 IGHAPVILSLTFNRKGNLFTSMDSIKIMD 124

RESULT 11  
 YLN2\_CAEEL STANDARD; PRT; 415 AA.  
 AC 018964;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 46.2 kDa Trp-Asp repeats containing protein D2013.2 in  
 DE chromosome II.  
 GN D2013.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peleodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mortimore B.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 FIVE-TYPE ZINC FINGER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z47808; CAA87770.1; -  
 DR Morneped: D2013.2; CE00928.  
 DR InterPro: IPR001680; WD40.  
 DR InterPro: IPR000306; znf\_FYVE.  
 DR Pfam: PF00400; WD40; 5.  
 DR Pfam: PF01363; FYVE; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR SMART: SM00064; FYVE; 1.  
 DR SMART: SM00320; WD40; 5.  
 DR PROSITE: PS50178; ZF\_FYVE; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.

DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein: Repeat; WD repeat; zinc-finger.  
 FT REPEAT 71 103 WD 1.  
 FT REPEAT 119 148 WD 2.  
 FT REPEAT 202 232 WD 3.  
 FT ZN\_FING 286 357 FYVE-TYPE.  
 FT REPEAT 373 403 WD 4.  
 SQ SEQUENCE 415 AA; 46239 MW; 8927612C040F5E43 CRC64;

Query Match 34.3%; Score 59; DB 1; Length 415;  
 Best Local Similarity 36.7%; Pred. No. 0.2; Mismatches 12; Indels 0; Gaps 0;  
 Matches 11; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

OY 2 IHHCVALHLRFN-NGMAYTCSKRSIAVMD 31  
 Db 374 IHTGTAMHLEDTGLVTSQNRVIMMD 403

RESULT 12  
 YDPL\_SCHPO STANDARD; PRT; 431 AA.  
 AC 014021;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical WD-repeat protein C29A4.18 in chromosome I.  
 DE SPAC29A4.18.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RX MEDLINE-21848401; PubMed-11859360;  
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Haidago J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,  
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerritielli L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shporkovskii G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe";  
 RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE HIRI FAMILY OF WD-REPEAT PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 Ansojge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
 Mewes H.-W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,  
 Wambolt R., Korn B., Klein M., Poustka A.;  
 "Towards a catalog of human genes and proteins: sequencing and  
 analysis of 500 novel complete protein coding human cDNAs.";  
 RL Genome Res. 11:422-435(2001).  
 [4]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Hypothalamus;  
 RX MEDLINE=20402571; PubMed=10931946;  
 RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,  
 Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,  
 Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,  
 Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,  
 Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;  
 "Gene expression profiling in the human hypothalamus-pituitary-adrenal  
 axis and full-length cDNA cloning.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).  
 [5]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=pituitary;  
 RA Song H., Peng Y., Dai M., Huang Q., Mao Y., Zhang Q., Mao M., Fu G.,  
 Luo M., Chen J., Hu R.;  
 Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Colon;  
 RA Strausberg R.;  
 Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
 produced by alternative splicing.  
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AF072880; AAD28808.1; -  
 DR EMBL: AF069313; AAD20954.2; -  
 DR EMBL: AF240696; AAF82746.1; -  
 DR EMBL: AL110243; CAB53693.1; -  
 DR EMBL: AF112205; AAF17193.1; -  
 DR EMBL: AF106683; AAD43036.1; -  
 DR EMBL: BC021110; AAH21110.1; -  
 DR InterPro: IPR001496; SOCS;  
 DR InterPro: IPR001680; WD40;  
 DR Pfam: PF00400; WD40; 6;  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 2;  
 DR SMART: SM00253; SOCS; 1;  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2;  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6;  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1;  
 DR PROSITE: PS50225; SOCS; 1;  
 KW Repeat; WD repeat; Alternative splicing.  
 FT REPEAT 32 71 WD 1.  
 FT REPEAT 124 165 WD 2.  
 FT REPEAT 168 208 WD 3.  
 FT REPEAT 212 251 WD 4.  
 FT REPEAT 254 293 WD 5.

FT REPEAT 309 346 WD 6.  
 FT DOMAIN 372 421 SOCS BOX.  
 FT VARSPLIT 14 159 MISSING (IN ISOFORM 2).  
 FT CONFLICT 16 16 L -> S (IN REF. 3 AND 4).  
 FT CONFLICT 178 178 F -> L (IN REF. 5).  
 FT CONFLICT 231 231 S -> P (IN REF. 5).  
 SQ SEQUENCE 421 AA; 47432 MW; 650d942e97d4bd CRC64;  
 Query Match 34.0%; Score 58.5; DB 1; Length 421;  
 Best Local Similarity 35.3%; Pred. No. 0.24;  
 Matches 12; Conservative 9; Mismatches 10; Indels 3; Gaps 1;  
 Oy 1 LIRHCEAVLHLEF--NNGMWTCSKDRSLAVWD 31  
 Db 166 LVDRHEVVRDLTFAPDGSLLVLSASRDKTLRWMD 199  
 RESULT 15  
 ID WSB1\_MOUSE STANDARD; PRT; 421 AA.  
 WS WSB1\_MOUSE  
 AC 054927;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE WD repeat and SOCS box containing protein 1 (WSB-1).  
 GN WSB1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIH Swiss;  
 RX MEDLINE=98081836; PubMed=9419338;  
 RA Hilton D.J., Richardson R.T., Alexander W.S., Viney E.M.,  
 Wilson T.A., Sprigg N.S., Starr R., Nicholson S.E., Meccalf D.,  
 Nicola N.A.;  
 "Twenty proteins containing a C-terminal SOCS box form five structural  
 classes.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 95:114-119(1998).  
 RL  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AF033186; AAB96647.1; -  
 DR InterPro: IPR001496; SOCS;  
 DR InterPro: IPR001680; WD40;  
 DR Pfam: PF00400; WD40; 6;  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 2;  
 DR SMART: SM00253; SOCS; 1;  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2;  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6;  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1;  
 DR PROSITE: PS50225; SOCS; 1;  
 KW Repeat; WD repeat.  
 FT REPEAT 124 165 WD 1.  
 FT REPEAT 168 208 WD 2.  
 FT REPEAT 212 251 WD 3.  
 FT REPEAT 254 293 WD 4.  
 FT REPEAT 309 346 WD 5.  
 FT DOMAIN 372 421 SOCS BOX.  
 SQ SEQUENCE 421 AA; 47065 MW; 4003d1ffe7a9a2ef CRC64;  
 Query Match 34.0%; Score 58.5; DB 1; Length 421;

Best Local Similarity 35.3%; Pred. No. 0.24;  
Matches 12; Conservative 9; Mismatches 10; Indels 3; Gaps 1;

Qy 1 LIHCEAVLHLRF--NNGMMVTCCKDRSIAVWD 31  
| | | | | : : : | | |  
Db 166 LVDHTEWVRDLTFAPDGSLLIVSASRDKTLRWVD 199

Search completed: April 10, 2003, 13:19:35  
Job time : 3.22777 secs



GenCore version 5.1.4-P5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 ; Search time 7.32521 Seconds  
(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_343\_373

Perfect score: 172  
Sequence: 1 LIHHCAYLHLRFNGMVTGSKDSIAVWD 31

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rident:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	172	100.0	122	13	08UUN3	08uun3 xenopus lae
2	172	100.0	510	5	044382	044382 drosophila
3	172	100.0	510	5	09YDE3	09yde3 drosophila
4	172	100.0	569	11	092159	092159 mus musculu
5	172	100.0	569	11	09Q015	09q015 mus musculu
6	167	97.1	569	11	09R1G7	09r1g7 mus musculu
7	157	91.3	430	5	09B454	09b454 heterodera
8	152	88.4	252	11	0922C7	0922c7 mus musculu
9	152	88.4	563	11	0923H0	0923h0 mus musculu
10	70	40.7	524	3	09P7C0	09p7c0 schizosacch
11	69	40.1	280	5	08STZ5	08stz5 encephalito
12	68	39.5	627	5	09VM11	09vm11 drosophila
13	67	39.0	442	3	096V17	096v17 pneumocysti
14	67	39.0	1609	5	09GRC3	09grc3 caenorhabdi
15	67	39.0	1893	5	017887	017887 caenorhabdi
16	66	38.4	331	5	019211	019211 caenorhabdi

17	66	38.4	436	5	09YAK0	09yak0 drosophila
18	65	37.8	573	11	09QXW2	09qxw2 mus musculu
19	64	37.2	310	4	075938	075938 homo sapien
20	64	37.2	357	4	095320	095320 homo sapien
21	64	37.2	357	4	096D17	096d17 homo sapien
22	64	37.2	370	11	09DCX3	09dcx3 mus musculu
23	64	37.2	414	5	08SM59	08sm59 encephalito
24	64	37.2	566	4	09H6A8	09h6a8 homo sapien
25	64	37.2	566	4	0969U6	0969u6 homo sapien
26	63.5	36.9	459	4	09NV62	09nv62 homo sapien
27	63.5	36.9	459	4	09HADA	09had4 homo sapien
28	63.5	36.9	942	5	096611	096611 dictyosteli
29	63	36.6	299	10	094AH2	094ah2 arabidopsis
30	63	36.6	343	10	09PG50	09pg50 arabidopsis
31	62.5	36.3	810	3	09PEV7	09pev7 neurospora
32	62.5	36.3	1326	5	09VZF4	09vzf4 drosophila
33	62	36.0	357	11	09DOK7	09dok7 mus musculu
34	62	36.0	360	4	09C069	09c069 homo sapien
35	62	36.0	360	4	096EE3	096ee3 homo sapien
36	62	36.0	360	4	096EE3	096ee3 homo sapien
37	62	36.0	421	4	096MH3	096mh3 homo sapien
38	62	36.0	507	3	09UT85	09ut85 schizosacch
39	62	36.0	553	4	09NUX6	09nux6 homo sapien
40	62	36.0	561	4	096R12	096r12 homo sapien
41	62	36.0	589	4	096LE0	096le0 homo sapien
42	62	36.0	627	4	096A16	096a16 homo sapien
43	62	36.0	629	11	08VHP4	08vhp4 mus musculu
44	62	36.0	629	11	08VBV4	08vbv4 mus musculu
45	62	36.0	707	4	0969H0	0969h0 homo sapien

## ALIGNMENTS

## RESULT 1

08UUN3 PRELIMINARY: PRT: 122 AA.

AC 08UUN3: 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Beta-TrCP protein (Fragment).  
GN BETA-TRCP.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Carnevali F.;  
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ballarino M.;  
RT "Analisi strutturale e funzionale del gene beta-TrCP in Xenopus laevis".  
RT Thesis (2001),  
RL Department of Genetics and Molecular Biology "Charles Darwin",  
RL University of Rome La Sapienza, Rome, Italy.  
DR EMBL: AJ428938; CAD21935.1; -  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 3.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR PRODOM: PD0000018; WD40; 2.  
DR SMART: SM00320; WD40; 3.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_2.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 3.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
FT NON\_TER 1  
FT 122  
FT SEQUENCE 122 AA; 13682 MW; 36BB6B39AC8F5387 CRC64;

Query Match 100.0%; Score 172; DB 13;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIHCEAVLHLRFNNGMMVTCSDRSTIAVMD 31  
 DB 13 LIHCEAVLHLRFNNGMMVTCSDRSTIAVMD 43

## RESULT 2

044382 PRELIMINARY; PRT; 510 AA.

DT 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

SLMB OR SLMB OR CG3412.  
 Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI\_TaxID=7227;

SEQUENCE FROM N.A.

RA MEDLINE=98121115; PubMed=9461217;

RA Jiang J., Struhl G.;

RT Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slmb."

RL Nature 391:493-496(1998).

CC -1- SIMILARITY: CONTAINS 7 WD\_REPEATS (TRP-ASP DOMAINS).

DR EMBL: AF032878; AAC38652.1;

DR InterPro: IPR001810; F-box.

DR Pfam: PF00646; F-box; 1.

DR PRINTS: PR00320; GPROTEINRPT.

DR PRODOM: PD000018; WD40; 4.

DR SMART: SM00256; FBOX; 1.

DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.

DR PROSITE: PS50082; WD\_REPEATS\_2; 7.

DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.

REPEAT: WD repeat.

SEQUENCE 510 AA; 58952 MM; DB024303730A58 CRC64;

Query Match

Best Local Similarity 100.0%; Score 172; DB 5; Length 510;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIHCEAVLHLRFNNGMMVTCSDRSTIAVMD 31

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.E., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos R.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bruttner P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunko B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kul D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacלב J.M.,

RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodgett J., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu D., Zhao Q., Zhao L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RA "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

RT [2]

RP SEQUENCE FROM N.A.

RC TISSUE-OVARY, AND IMAGINAL DISC;

RA MEDLINE=20245299; PubMed=10781936;

RA Miletich I., Limbourg-Bouchon B.;

RT "Drosophila null slmb clones transiently deregulate Hedgehog-

RT independent transcription of wingless in all limb discs, and induce

RT decapentaplegic transcription linked to imaginal disc regeneration.";

RL Mech. Dev. 93:15-26(2000).

CC -1- SIMILARITY: CONTAINS 7 WD\_REPEATS (TRP-ASP DOMAINS).

DR EMBL: AE003733; AAF55853.1;

DR EMBL: AF222924; AAF63214.1;

DR EMBL: AF222923; AAF63213.1;

DR FLYBASE: FBgn0023423; slmb.

DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.

DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 7.

DR PRINTS: PR00320; GPROTEINRPT.

DR PRODOM: PD000018; WD40; 4.

DR SMART: SM00256; FBOX; 1.

DR SMART: SM00320; WD40; 7.

DR PROSITE: PS50181; FBOX; 1.

DR PROSITE: PS00678; WD\_REPEATS\_1; 5.

DR PROSITE: PS50082; WD\_REPEATS\_2; 7.

DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.

REPEAT: WD repeat.

SEQUENCE 510 AA; F4D5DF126F58A012 CRC64;

Query Match

Best Local Similarity 100.0%; Score 172; DB 5; Length 510;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIHCEAVLHLRFNNGMMVTCSDRSTIAVMD 31

Db 286 LIHCEAVLHLRFNNGMVTCSKDRSLAVWD 316

## RESULT 4

Q92159 PRELIMINARY: PRT: 569 AA.

AC 092159; (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 20, Last annotation update)  
 DE Beta-transducin repeat containing protein.  
 GN BTRC.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=99145465; PubMed=9990853;  
 RA Spencer E., Jiang J., Chen Z.J.;  
 RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein  
 SLIMB/beta-TrCP.";  
 RL Genes Dev. 13:284-294(1999).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF112979; AAD04181.1; -  
 DR MGI: 1338871; Btrc.

DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.

DR PROSITE: PS0181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;

Query Match 100.0%; Score 172; DB 11; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 2,6e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHCEAVLHLRFNNGMVTCSKDRSLAVWD 31  
 Db 343 LIHCEAVLHLRFNNGMVTCSKDRSLAVWD 373

## RESULT 5

Q9Q015 PRELIMINARY: PRT: 569 AA.

AC 09Q015; (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 20, Last annotation update)  
 DE Ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)  
 DE (F-box-WD40 repeat protein 1).  
 GN BTRC OR FBXW1.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=99199275; PubMed=10097128;  
 RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,  
 RA Hattori K., Higashi H., Nakano H., Okumura K., Ono K., Good R.A.,  
 RA Nakayama K.-I.;  
 RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a  
 RT ubiquitin ligase Skp1/Cul 1/F-box protein FWD1.";

Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).

[2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99075339; PubMed=9859996;  
 RA Yaron A., Hatzudai A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
 RT "Identification of the receptor component of the IkappaBalpha-  
 RT ubiquitin ligase.";  
 RL Nature 396:590-594(1998).

[3]  
 RP SEQUENCE FROM N.A.

RA Strausberg R.;  
 RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

[4]  
 RP SEQUENCE FROM N.A.

RC STRAIN-129/SV;  
 RX MEDLINE=21601157; PubMed=11735228;  
 RA Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,  
 RA Nakayama K.-I.;

RT "Characterization of a Mouse Gene (fbxw6) That Encodes a Homologue of  
 RT Caenorhabditis elegans SEL-10.";  
 RL Genomics 78:214-222(2001).

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF081887; AAD17755.1; -  
 DR EMBL: AF099932; AAD08701.1; -  
 DR EMBL: BC003989; AAH03989.1; -  
 DR EMBL: AF391190; AAL40929.1; -  
 DR EMBL: AF391178; AAL40929.1; JOINED.  
 DR EMBL: AF391179; AAL40929.1; JOINED.  
 DR EMBL: AF391180; AAL40929.1; JOINED.  
 DR EMBL: AF391181; AAL40929.1; JOINED.  
 DR EMBL: AF391182; AAL40929.1; JOINED.  
 DR EMBL: AF391183; AAL40929.1; JOINED.  
 DR EMBL: AF391184; AAL40929.1; JOINED.  
 DR EMBL: AF391185; AAL40929.1; JOINED.  
 DR EMBL: AF391186; AAL40929.1; JOINED.  
 DR EMBL: AF391187; AAL40929.1; JOINED.  
 DR EMBL: AF391188; AAL40929.1; JOINED.  
 DR EMBL: AF391189; AAL40929.1; JOINED.  
 DR MGI: 1338871; Btrc.

DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.

DR PROSITE: PS0181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.

QY 1 LIHCEAVLHLRFNNGMVTCSKDRSLAVWD 31  
 Db 343 LIHCEAVLHLRFNNGMVTCSKDRSLAVWD 373

Query Match 100.0%; Score 172; DB 11; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 2,6e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIHCEAVLHLRFNNGMVTCSKDRSLAVWD 31  
 Db 343 LIHCEAVLHLRFNNGMVTCSKDRSLAVWD 373

## RESULT 6

Q9R1G7 PRELIMINARY: PRT: 569 AA.

AC 09R1G7; (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 21, Last annotation update)  
 DE Beta-transducin repeat-containing protein.  
 GN BTRC.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Winston J., Ellledge S.J., Harper J.W.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF110396; AAD41023.1; -.  
 DR MGD: MGI:138871; Birc.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR Prodom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

Query Match 97.1%; Score 167; DB 11; Length 569;  
 Best Local Similarity 96.8%; Pred. No. 1.6e-17;  
 Matches 30; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LHHCEAVLHLRFNNGMVTCSKDRSIAYVD 31  
 Db 343 LHHCEAVLHLRFNNGMVTCSKDRSIAYVD 373

RESULT 7  
 Q9BJ54 PRELIMINARY; PRT; 430 AA.  
 AC Q9BJ54;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Beta-transducin repeat-containing protein (Fragment).  
 OS Heterodera glycines (Soybean cyst nematode).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchima;  
 OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.  
 NX NCBI\_TaxID=51029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kovaleva E.S., Yakovlev A.G., Master E.P.;  
 RT "Plant parasitic nematode b-TRCP.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF339101; AAK26376.1; -.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR Prodom: PD000018; WD40; 4.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER  
 SQ SEQUENCE 430 AA; 47916 MW; 4ABC3FFD2FE3A50B CRC64;

Query Match 91.3%; Score 157; DB 5; Length 430;  
 Best Local Similarity 87.1%; Pred. No. 4.2e-16;  
 Matches 27; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LHHCEAVLHLRFNNGMVTCSKDRSIAYVD 31  
 Db 167 LVHCEAVLHLRFNNGMVTCSKDRSIAYVD 197

RESULT 8  
 Q922C7 PRELIMINARY; PRT; 252 AA.  
 AC Q922C7;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Similar to f-box and WD-40 domain protein 1B (Fragment).  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: BC008552; AAH08552.1; -.  
 DR MGD: MGI:2144023; FBXW1B.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR Prodom: PD000018; WD40; 3.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER  
 SQ SEQUENCE 252 AA; 28424 MW; F71737C8D7A9F75F CRC64;

Query Match 88.4%; Score 152; DB 11; Length 252;  
 Best Local Similarity 90.3%; Pred. No. 1.4e-15;  
 Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LHHCEAVLHLRFNNGMVTCSKDRSIAYVD 31  
 Db 26 LHHCEAVLHLRFNNGMVTCSKDRSIAYVD 56

RESULT 9  
 Q923H0 PRELIMINARY; PRT; 563 AA.  
 AC Q923H0;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE F-box/WD40 repeat-containing protein HOS.  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bhattacharya N., Herter J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;  
 RT "Mouse homolog of HOS (mHOS) is overexpressed in skin tumors and  
 implicated in constitutive activation of NF-kappaB.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AY038079; AAK72095.1; -.  
 DR MGD: MGI:2144023; FBXW1B.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR Prodom: PD000018; WD40; 4.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.

RN [2]  
RP SEQUENCE FROM N. A.  
RC STRAIN-GB-M1;

Wang Z., Wu D., Woodage T., Morley K. C., Wu D., Yang S., Yao Q. A.,  
 RA Williams S. M., Woodage T., Morley K. C., Zhang G., Zhao Q., Zheng L.,  
 Ye J., Yeh R. F., Zaveri J. S., Zhan M.

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*";  
 RL Science 287:2185-2195(2000).  
 CC -1- SIMILARITY: CONTAINS 3 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AE003612; AAF52336.1; -;  
 DR FLYBASE: FBgn0031773; CG9144.  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF00646; F-box; 1.  
 DR SMART: SM00400; WD40; 3.  
 DR SMART: SM00256; F-box; 1.  
 DR SMART: SM00320; WD40; 2.  
 DR PROSITE: PS0181; F-box; 1.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 2.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 627 AA; 71908 MW; 11A2E2FEF4DE433 CRC64;

Query Match 39.5%; Score 68; DB 5; Length 627;  
 Best Local Similarity 43.3%; Pred. No. 0.046;  
 Matches 13; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

OY 4 HCEAVLHFR--NGMMVTCSDRSIAVWD 31  
 DB 128 HTHQVLHVSFAHNGEMFATCSKDGVIITWN 157

## RESULT 13

ID 096V17 PRELIMINARY; PRT; 442 AA.  
 AC 096V17;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Probable nuclear migration protein.  
 OS *Pneumocystis carinii*.  
 OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;  
 OC Pneumocystis.  
 OC NCBI\_TaxID=4754;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PNEUMOCYSTIS CARINII F. SP. CARINII;  
 RA Murphy L., Quail M., Harris D., Hall N., Wakefield A., Smolian A.G.,  
 RA Cushion M.T., Stringer J.R., Keely S.P., Bartell B.G.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AL592382; CAC43453.1; -;  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRODOM: PD000018; WD40; 2.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 5.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 442 AA; 49914 MW; 0B16704A217BC123 CRC64;

Query Match 39.0%; Score 67; DB 3; Length 442;  
 Best Local Similarity 40.6%; Pred. No. 0.045;  
 Matches 13; Conservative 7; Mismatches 8; Indels 4; Gaps 1;

OY 4 HCEAVLHFR--NGMMVTCSDRSIAVWD 31  
 DB 156 HTKAVLDIEFTFSDDLVSCLVSCSDLTIKIMD 187

## RESULT 14

ID 09GRC3 PRELIMINARY; PRT; 1609 AA.  
 AC 09GRC3;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE kinesin like protein KLP-12.  
 GN KLP-12.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA All M.Y., Siddiqui S.S.;  
 RT "C. elegans kinesin like protein, KLP-12";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AB035591; BAB16763.1; -;  
 DR HSP: P17119; 3KAR.  
 DR InterPro: IPR001752; kinesin\_motor.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00225; kinesin; 1.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PR00320; GPROTEINBPT.  
 DR PRINTS: PR00380; KINESINHEAVY.  
 DR PRODOM: PD000018; WD40; 1.  
 DR SMART: SM00129; KISC; 1.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PS00411; KINESIN\_MOTOR\_DOMAIN1; UNKNOWN\_1.  
 DR PROSITE: PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 DR PROSITE: PS500678; WD\_REPEATS\_1; UNKNOWN\_2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW ATP-binding; Coiled coil; Microtubules; Motor protein; Repeat;  
 KW WD repeat.  
 SQ SEQUENCE 1609 AA; 179646 MW; 5F710B840B4E9416 CRC64;

Query Match 39.0%; Score 67; DB 5; Length 1609;  
 Best Local Similarity 46.4%; Pred. No. 0.18;  
 Matches 13; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 4 HCEAVLHFR--NGMMVTCSDRSIAVWD 31  
 DB 1287 HARGVLSVDVNEKLVTCSDKRTAKIMD 1314

## RESULT 15

ID 017887 PRELIMINARY; PRT; 1893 AA.  
 AC 017887; 018007;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE T01G1.1 protein.  
 DE T01G1.1 protein.  
 GN T01G1.1.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kershaw J.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latteille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaison N., Smith A., Sonnenhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Wellstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*.";

RL Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA White S., Mortimore B.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: 292811; CAB07273.1; -  
 DR EMBL: 282271; CAB07273.1; JOINED.  
 DR EMBL: 282271; CAB05214.1; -  
 DR EMBL: 292811; CAB05214.1; JOINED.  
 DR HSSP: P17119; 3KAR.  
 DR InterPro: IPR001752; kinesin\_motor.  
 DR InterPro: IPR001441; UPP\_synth.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00225; kinesin; 1.  
 DR Pfam: PF01255; UPP\_synthetase; 1.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PR00320; GPOTEINRPT.  
 DR PRINTS: PR00380; KINESINHEAVY.  
 DR ProDom: PD000018; WD40; 1.  
 DR ProDom: PD003461; UPP\_synth; 1.  
 DR SMART: SM00129; KISC; 1.  
 DR SMART: SM00320; WD40; 3.  
 DR TIGRFAMs: TIGR00055; upps; 1.  
 DR PROSITE: PS00411; KINESIN\_MOTOR\_DOMAIN1; UNKNOWN\_1.  
 DR PROSITE: PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR ATP-binding; Coiled coil; Microtubules; Motor protein; Repeat;  
 KW WD repeat.  
 SQ SEQUENCE 1893 AA; 212825 MW; 166B7477C1BD810C CRC64;

Query Match 39.0%; Score 67; DB 5; Length 1893;  
 Best Local Similarity 46.4%; Pred. No. 0.22; Indels 0; Gaps 0;  
 Matches 13; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
 QY 4 HCEAVLHLRFNGMVTGSKDRSIAVWD 31  
 Db 1571 HARGVLSVDVNEKIMVTGSKDRITAKLWD 1598

Search completed: April 11, 2003, 11:51:36  
 Job time : 9.32521 secs

1. The first part of the report is a general introduction to the subject of the study. It discusses the importance of the study and the objectives of the research.

2.

3.



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:20:44 ; Search time 8.93666 Seconds  
(without alignments)  
432.406 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_387\_415

Petfect score: 151  
Sequence: 1 VGHRAAVNVDFDCKITVSASGDKRTIKVM 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A.Geneseq.101002:\*  
2: /SID2/gcgdata/geneseq/genesep-emb1/AA1980.DAT.\*  
3: /SID2/gcgdata/geneseq/genesep-emb1/AA1981.DAT.\*  
4: /SID2/gcgdata/geneseq/genesep-emb1/AA1982.DAT.\*  
5: /SID2/gcgdata/geneseq/genesep-emb1/AA1983.DAT.\*  
6: /SID2/gcgdata/geneseq/genesep-emb1/AA1984.DAT.\*  
7: /SID2/gcgdata/geneseq/genesep-emb1/AA1985.DAT.\*  
8: /SID2/gcgdata/geneseq/genesep-emb1/AA1986.DAT.\*  
9: /SID2/gcgdata/geneseq/genesep-emb1/AA1987.DAT.\*  
10: /SID2/gcgdata/geneseq/genesep-emb1/AA1988.DAT.\*  
11: /SID2/gcgdata/geneseq/genesep-emb1/AA1989.DAT.\*  
12: /SID2/gcgdata/geneseq/genesep-emb1/AA1990.DAT.\*  
13: /SID2/gcgdata/geneseq/genesep-emb1/AA1991.DAT.\*  
14: /SID2/gcgdata/geneseq/genesep-emb1/AA1992.DAT.\*  
15: /SID2/gcgdata/geneseq/genesep-emb1/AA1993.DAT.\*  
16: /SID2/gcgdata/geneseq/genesep-emb1/AA1994.DAT.\*  
17: /SID2/gcgdata/geneseq/genesep-emb1/AA1995.DAT.\*  
18: /SID2/gcgdata/geneseq/genesep-emb1/AA1996.DAT.\*  
19: /SID2/gcgdata/geneseq/genesep-emb1/AA1997.DAT.\*  
20: /SID2/gcgdata/geneseq/genesep-emb1/AA1998.DAT.\*  
21: /SID2/gcgdata/geneseq/genesep-emb1/AA1999.DAT.\*  
22: /SID2/gcgdata/geneseq/genesep-emb1/AA2000.DAT.\*  
23: /SID2/gcgdata/geneseq/genesep-emb1/AA2001.DAT.\*  
24: /SID2/gcgdata/geneseq/genesep-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	151	100.0	517	16 AAR85852	WD-40 domain-contg
2	151	100.0	542	21 AAY96696	Human E3 ubiquitin
3	151	100.0	542	22 AAM79127	Human protein SFO
4	151	100.0	542	22 AAM40208	Human polypeptide
5	151	100.0	550	22 AAM41994	Human polypeptide
6	151	100.0	569	20 AAY24054	A human beta-trans
7	151	100.0	569	21 AAB12812	Mouse ubiquitin 11
8	151	100.0	569	21 AAB12813	Human beta-transdu
9	151	100.0	569	21 AAY96697	Human beta-TrCP.
10	151	100.0	569	21 AAY83041	F-box protein FBP-

11	151	100.0	569	21 AAY83250	F-box protein hbet
12	151	100.0	569	21 AAY83254	F-box protein FMD1
13	151	100.0	569	21 AAY44249	Human cell signal
14	151	100.0	569	22 AAB48298	Human zell protein
15	151	100.0	579	22 AAB78583	Human protein SFO
16	151	100.0	590	22 AAM00847	Human bone marrow
17	151	100.0	605	22 AAM78582	Human protein SFO
18	151	100.0	608	22 AAM00960	Human bone marrow
19	151	100.0	632	22 AAM78584	Human protein SFO
20	151	100.0	654	22 AAM79566	Human protein SFO
21	151	100.0	654	22 AAM79567	Human protein SFO
22	151	100.0	654	22 AAM79568	Human protein SFO
23	151	100.0	654	22 AAM79569	Human protein SFO
24	147	97.4	510	22 AAB84938	Human protein SFO
25	82	54.3	17	21 AAY96714	Human protein SFO
26	78	51.7	1326	22 AAB67237	Human protein SFO
27	78	51.7	1326	22 AAB67238	Human protein SFO
28	78	51.7	1326	22 AAB67239	Human protein SFO
29	76	50.3	514	16 AAR85881	Human protein SFO
30	75	49.7	513	22 AAM40282	Human protein SFO
31	75	49.7	513	22 AAB93194	Human protein SFO
32	75	49.7	513	22 AAB85110	Human protein SFO
33	75	49.7	525	22 AAM42068	Human protein SFO
34	75	49.7	587	20 AAY03204	Human protein SFO
35	75	49.7	860	22 AAB62502	Human protein SFO
36	74	49.0	33	16 AAB84944	Human protein SFO
37	74	49.0	690	21 AAY49948	Human protein SFO
38	74	49.0	744	21 AAB56769	Human protein SFO
39	74	49.0	906	16 AAB85853	Human protein SFO
40	73	48.3	234	21 AAG23335	Human protein SFO
41	73	48.3	244	21 AAG23334	Human protein SFO
42	73	48.3	310	21 AAG28461	Human protein SFO
43	73	48.3	310	21 AAG28462	Human protein SFO
44	73	48.3	311	21 AAG04819	Human protein SFO
45	73	48.3	316	21 AAG28460	Human protein SFO

## ALIGNMENTS

RESULT 1  
ID AAR85852 standard: peptide: 517 AA.  
XX AAR85852:  
XX 13-SEP-1996 (first entry)  
XX WD-40 domain-contg. beta-TRCP protein.  
XX WD40 repeat region: beta-transducin: protein-protein interaction: drug;  
XX intracellular signalling: protein kinase C; homology; motif: modulator;  
XX receptors of activated protein kinase; enzyme activity: isozyme; human.  
XX Synthetic.  
XX W09521252-A2.  
XX 10-AUG-1995.  
XX 31-JAN-1995: 95WO-US01210.  
XX 01-FEB-1994: 94US-0190802.  
XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
XX Mochly-Rosen D, Ron D:  
XX WPI: 1995-283772/37.  
XX New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the  
XX activity of a protein, eg. protein kinase C, which interacts with a  
XX protein contg. a WD-40 region.

XX Example 5; Page 80-82; 351pp; English.  
PS  
XX  
CC Proteins AAR85851-92 are protein which contain at least one WD-40 (also  
CC called beta-transducing homologous) amino acid repeat motifs. The WD-40  
CC regions are involved in protein-protein interactions between proteins  
CC involved in intracellular signalling. An example of such an interaction  
CC is between protein kinase C and receptors of activated protein kinase  
CC (RACK), esp. RACK-1 (AAR85850). Proteins AAR85851-82 were isolated based  
CC on homology with beta-transducin, whereas proteins AAR85882-92 were  
CC isolated based on homology with the WD-40 consensus sequence (AAR85893).  
CC The proteins were used to construct the peptides AAR84928-R85063 and  
CC AAR85786-R85842. The peptides can be used to identify target proteins  
CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of  
CC proteins involved in protein-protein interaction and to screen for drugs  
CC that will affect protein-protein interaction involving WD-40 domains.  
XX  
SQ Sequence 517 AA;  
Query Match 100.0%; Score 151; DB 16; Length 517;  
Best Local Similarity 100.0%; Pred. No. 2,9e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VGHRAAVNVVDDPKYIVASGDRITIKYW 29  
DB 351 VGHRAAVNVVDDPKYIVASGDRITIKYW 379  
RESULT 2  
AAY96696  
ID AAY96696 standard; Protein: 542 AA.  
AC AAY96696;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Human E3 ubiquitin ligase.  
XX  
DE E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
KW anti-inflammatory; immunosuppressive; cytosolic.  
XX  
OS Homo sapiens.  
XX  
PN WO200034447-A2.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99WO-US29371.  
XX  
PR 10-DEC-1998; 98US-0210060.  
XX  
PA (SIGMA) SIGMAL PHARM INC.  
XX (YISS) YISSUM RES & DEV CO.  
XX  
PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;  
PI Layon I, Yaron A;  
XX  
DR WPI: 2000-431294/37.  
DR N-PSDB: AAK51228.  
XX  
PT Polypeptide enhancing phosphorylated I-kappa-B ubiquitination useful for  
PT treating disorder associated with NF-kappa-B activation e.g. cancer,  
PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
PT variant  
XX  
PS Claim 1; Page 70-72; 77pp; English.  
XX  
CC This is human E3 ubiquitin ligase (E3), which is homologous to human  
CC beta-TrCP, an F-box/WD protein family member. E3 enhances ubiquitination  
CC of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor  
CC kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the  
CC ubiquitin pathway is useful for identifying modulators of this process

CC for use in treating diseases associated with activation of NF-kappa-B. In  
CC vitro analysis suggests that deletion of the F-box results in a protein  
CC that functions as a dominant negative molecule in vivo. Transient  
CC over-expression of delta-beta-TrCP (a deletion mutant) inhibited the  
CC degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells,  
CC resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be  
CC used to screen for modulators of NF-kappa-B activity. E3 and beta-TrCP  
CC can be used to modulate NF-kappa-B to treat inflammatory diseases,  
CC autoimmune diseases, cancer and viral infections.  
XX  
SQ Sequence 542 AA;  
Query Match 100.0%; Score 151; DB 21; Length 542;  
Best Local Similarity 100.0%; Pred. No. 3.1e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VGHRAAVNVVDDPKYIVASGDRITIKYW 29  
DB 360 VGHRAAVNVVDDPKYIVASGDRITIKYW 388  
RESULT 3  
AAM79127  
ID AAM79127 standard; Protein: 542 AA.  
AC AAM79127;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human protein SEQ ID NO 1789.  
XX  
DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200157190-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US04098.  
XX  
PR 03-FEB-2000; 2000US-0496914.  
XX  
PR 27-APR-2000; 2000US-0560875.  
XX  
PR 20-JUN-2000; 2000US-0598075.  
XX  
PR 19-JUL-2000; 2000US-0620325.  
XX  
PR 01-SEP-2000; 2000US-0654936.  
XX  
PR 15-SEP-2000; 2000US-0663561.  
XX  
PR 20-OCT-2000; 2000US-0693325.  
XX  
PR 30-NOV-2000; 2000US-0728422.  
XX  
PA (HYSEQ) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Dymnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
DR WPI: 2001-476283/51.  
DR N-PSDB: AAK52260.  
XX  
PT Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
PS Claim 20; Page 4147-4148; 6221pp; English.  
XX  
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM0020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX Sequence 542 AA;

Query Match 100.0%; Score 151; DB 22; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVDPDDKYIVSASGDRITIKW 29  
 |||||  
 Db 360 VGHRAAVNVDPDDKYIVSASGDRITIKW 388

RESULT 4  
 AAM40208  
 ID AAM40208 standard; Protein; 542 AA.

XX AC AAM40208;  
 XX DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 3353.

XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX DR WPI: 2001-442253/47.  
 XX DR N-PSDB: AAI59364.

PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Example 5; SEQ ID NO 3353; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemia and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 542 AA;

Query Match 100.0%; Score 151; DB 22; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVDPDDKYIVSASGDRITIKW 29  
 |||||  
 Db 360 VGHRAAVNVDPDDKYIVSASGDRITIKW 388

RESULT 5  
 AAM41994  
 ID AAM41994 standard; Protein; 550 AA.

XX AC AAM41994;

XX DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6925.

XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX DR WPI: 2001-442253/47.  
 XX DR N-PSDB: AAI61150.

PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6925; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,



DB N-PSDB; AAA73131.  
 XX F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of IkappaB or beta-catenin  
 PS Claim 2; Page 9-10; 19pp; Japanese.  
 XX  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 CC  
 SQ Sequence 569 AA:  
 QY  
 Query Match 100.0%; Score 151; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 VGHRAAVNVVDFDDKTYIVSASGDRITIKW 29  
 ||||||||||||||||||||||||||||  
 DB 387 VGHRAAVNVVDFDDKTYIVSASGDRITIKW 415

RESULT 8  
 AAB12813  
 ID AAB12813 standard; protein; 569 AA.  
 AC AAB12813;  
 XX  
 DT 27-NOV-2000 (first entry)  
 XX  
 DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.  
 XX  
 KW ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;  
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KW beta-TrCP.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP2000166542-A.  
 XX  
 PD 20-JUN-2000.  
 XX  
 PF 02-DEC-1998; 98JP-0343437.  
 XX  
 PR 02-DEC-1998; 98JP-0343437.  
 XX  
 PA (KAGAKU) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 XX  
 DR WPI: 2000-485550/43.  
 XX  
 DR N-PSDB; AAA73132.  
 PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of IkappaB or beta-catenin  
 PS Claim 3; Page 10-12; 19pp; Japanese.  
 XX  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 CC

SQ Sequence 569 AA:  
 QY  
 Query Match 100.0%; Score 151; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 VGHRAAVNVVDFDDKTYIVSASGDRITIKW 29  
 ||||||||||||||||||||||||||||  
 DB 387 VGHRAAVNVVDFDDKTYIVSASGDRITIKW 415

RESULT 9  
 AAY96697  
 ID AAY96697 standard; protein; 569 AA.  
 AC AAY96697;  
 XX  
 DT 26-SEP-2000 (first entry)  
 XX  
 DE Human beta-TrCP.  
 XX  
 KW E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KW anti-inflammatory; immunosuppressive; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200034447-A2.  
 XX  
 PD 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99WO-US29371.  
 XX  
 PR 10-DEC-1998; 98US-0210060.  
 XX  
 PA (SIGMA) SIGNAL PHARM INC.  
 PA (YISS) YISSUM RES & DEV CO.  
 XX  
 PI Manning AM, Mercurio F, Amlt S, Ben-eriah Y, Davis M, Hatzubai A;  
 PI Lavon I, Yaron A;  
 XX  
 DR WPI: 2000-431294/37.  
 XX  
 DR N-PSDB; AAA51229.  
 PT polypeptide enhancing phosphorylated IkappaB ubiquitination useful for  
 PT treating disorder associated with NF-kappaB activation e.g. cancer,  
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
 PT variant  
 PS Claim 21; Page 72-74; 77pp; English.  
 XX  
 CC Human beta-TrCP, an F-box/WD protein family member, has been shown to  
 CC have homology to human E3 ubiquitin ligase (E3). E3 enhances  
 CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of  
 CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B  
 CC degradation via the ubiquitin pathway is useful for identifying  
 CC modulators of this process for use in treating diseases associated with  
 CC activation of NF-kappa-B. In vitro analysis suggests that deletion of  
 CC the F-box results in a protein that functions as a dominant negative  
 CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a  
 CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha  
 CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated  
 CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B  
 CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat  
 CC inflammatory diseases, autoimmune diseases, cancer and viral infections.  
 CC  
 SQ Sequence 569 AA:  
 QY  
 Query Match 100.0%; Score 151; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 VGHRAAVNVVDFDDKTYIVSASGDRITIKW 29

Db 387 VGHRAAVNVDFDDKTYVSASGDRITIKW 415

# RESULT 10

AAV83041

ID AAV83041 standard; Protein; 569 AA.

AC AAV83041;

DT 16-AUG-2000 (first entry)

DE F-box protein FBP-1.

OS F-box protein; FBP; diagnosis; treatment; screening; agonist; antagonist; proliferative disorder; differentiative disorder; breast cancer; prostate cancer; ovarian cancer; cancer;

KW small cell lung carcinoma; immune disorder; cardiovascular disorder; inflammatory disorder; human.

XX Homo sapiens.

PN WO200012679-A1.

PD 09-MAR-2000.

PF 27-AUG-1999; 99WO-0519560.

PR 28-AUG-1998; 98US-0098355.

PR 03-FEB-1999; 99US-0118568.

PR 15-MAR-1999; 99US-0124449.

PA (UYNV ) UNITV NEW YORK STATE.

PI Chiur DS, Pagano M, Latres E;

DR WPI; 2000-256635/22.

DR N-PSDB; AA293350.

PT Novel nucleic acid for screening compounds useful for treating proliferative and differentiative disorders such as cancer and immune disorders comprises sequences encoding ubiquitin ligases.

PS Disclousure; Figure 3a; 245pp; English.

CC Nucleic acids encoding substrate-targeting subunits of ubiquitin ligases with F-box motifs (F-box proteins) are useful for diagnosis of proliferative and differentiated related disorders by measuring FBP gene expression. Cells expressing such proteins or their fragments are useful for screening compounds. The compounds are agonists or antagonists, which are useful for treating a breast, ovarian and prostate cancer and small cell lung carcinoma and also major opportunistic infections, immune disorders, cardiovascular diseases and inflammatory disorders. FBP protein, analogs, derivatives and their subsequences, anti-FBP antibodies are also useful in diagnosis of the disorders.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

AC AAV83250;

DT 16-AUG-2000 (first entry)

DE F-box protein hBetatTCP.

KW ubiquitin ligase; SCF; F-box protein; targeted degradation; destabilisation; proteolysis; drug discovery; gene therapy; cancer; oncoprotein; Huntington's disease; gene knockout; delivery systems; human.

XX Homo sapiens.

PN WO200022110-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-0523705.

PR 09-OCT-1998; 98US-0103787.

PA (HARD ) HARVARD COLLEGE.

PI Zhou P, Howley P;

DR WPI; 2000-317970/27.

DR N-PSDB; AA293710.

PT Targeting degradation of polypeptide useful for treating cancer and other proliferative disorders. Involves conjugating polypeptide with ubiquitin protein ligase or inhibiting ubiquitination using organic compound

PS Claim 9; Page 171; 185pp; English.

XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin ligases) which can be used for the targeted degradation of a target polypeptide in vivo. Targeted degradation is achieved by expressing the ubiquitin ligase in a cell linked to the interaction domain of the target polypeptide and thereby recruiting the target polypeptide to the ubiquitin ligase. Such methods are useful for decreasing or increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating diseases caused by the presence of abnormal amount of the specific polypeptides. For drug discovery and for gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

AC AAV83254;

DT 16-AUG-2000 (first entry)

DE F-box protein FWDJP.

KW ubiquitin ligase; SCF; F-box protein; targeted degradation; destabilisation; proteolysis; drug discovery; gene therapy; cancer; oncoprotein; Huntington's disease; gene knockout; delivery systems; human.

XX Homo sapiens.

PN WO200022110-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-0523705.

PR 09-OCT-1998; 98US-0103787.

PA (HARD ) HARVARD COLLEGE.

PI Zhou P, Howley P;

DR WPI; 2000-317970/27.

DR N-PSDB; AA293710.

PT Targeting degradation of polypeptide useful for treating cancer and other proliferative disorders. Involves conjugating polypeptide with ubiquitin protein ligase or inhibiting ubiquitination using organic compound

PS Claim 9; Page 171; 185pp; English.

XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin ligases) which can be used for the targeted degradation of a target polypeptide in vivo. Targeted degradation is achieved by expressing the ubiquitin ligase in a cell linked to the interaction domain of the target polypeptide and thereby recruiting the target polypeptide to the ubiquitin ligase. Such methods are useful for decreasing or increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating diseases caused by the presence of abnormal amount of the specific polypeptides. For drug discovery and for gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC



PT Human cell signaling proteins useful for, e.g. diagnosing cell  
 PT proliferative and inflammatory disorders  
 XX  
 PS Claim 1; Page 77-78; 90pp; English.  
 XX  
 CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded  
 CC by cDNA obtained from Incyte clone 3239149 of COLACTO1 library. It is  
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is  
 CC found to be homologous to beta transducin repeats containing  
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as  
 CC hybridisation probe for detecting CSIGP related sequences or allelic  
 CC variants. Recombinant CSIGP can be produced in host cells by transforming  
 CC with genetically engineered vectors. Agonists or antagonists can be  
 CC used in the treatment of cell proliferative and inflammatory disorders  
 CC associated with decreased or increased CSIGP expression. CSIGP is used in  
 CC the diagnosis, prevention and treatment of cell proliferative disorders  
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.  
 XX  
 SQ Sequence 569 AA:  
 Query Match 100.0%; Score 151; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VGHRAAVNVDPDKYIVSASGDRITKW 29  
 DB 387 VGHRAAVNVDPDKYIVSASGDRITKW 415  
 RESULT 14  
 AAB48298  
 ID AAB48298 standard; protein; 569 AA.  
 AC AAB48298;  
 XX  
 XX 02-APR-2001 (first entry)  
 DT  
 XX  
 DE Human ZF11 protein.  
 XX  
 KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;  
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
 KW Bad; Bcl-2; tumour; cytostatic.  
 OS Homo sapiens.  
 XX  
 XX MO200075184-A1.  
 PN  
 XX 14-DEC-2000.  
 PD  
 XX 05-JUN-2000; 2000MO-US15449.  
 PF  
 XX 04-JUN-1999; 99US-0137494.  
 PR  
 XX (UYVA) UNIV YALE.  
 PA  
 XX Zhang H, Tsvetkov LM, Kondo T;  
 PI WPI: 2001-061703/07.  
 DR N-PSDB: AAC84610.  
 DR  
 XX Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
 PT involves altering levels of proteins such as S-phase kinase associated  
 PT proteins 1, 2 and cullin/CDC53 proteins -  
 PS Claim 3; Page 130-132; 162pp; English.  
 XX  
 CC The invention relates to methods of altering the polypeptide levels in a  
 CC cell, using proteins selected from S-phase kinase associated proteins 1  
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
 CC cullin/CDC53 family of proteins). The method is useful for altering the  
 CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for

CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
 CC that modulate interactions between SKP and target proteins are useful for  
 CC treating tumours.  
 CC  
 XX  
 SQ Sequence 569 AA:  
 Query Match 100.0%; Score 151; DB 22; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VGHRAAVNVDPDKYIVSASGDRITKW 29  
 DB 387 VGHRAAVNVDPDKYIVSASGDRITKW 415  
 RESULT 15  
 AAM78583  
 ID AAM78583 standard; Protein; 579 AA.  
 AC AAM78583;  
 XX  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX  
 DE Human protein SEQ ID NO 1245.  
 XX  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200157190-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 05-FEB-2001; 2001WO-US04098.  
 PF  
 XX 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R.  
 XX WPI: 2001-476283/51.  
 DR N-PSDB: AAK51716.  
 DR  
 XX Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 PT Claim 20; Page 3504-3505; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666



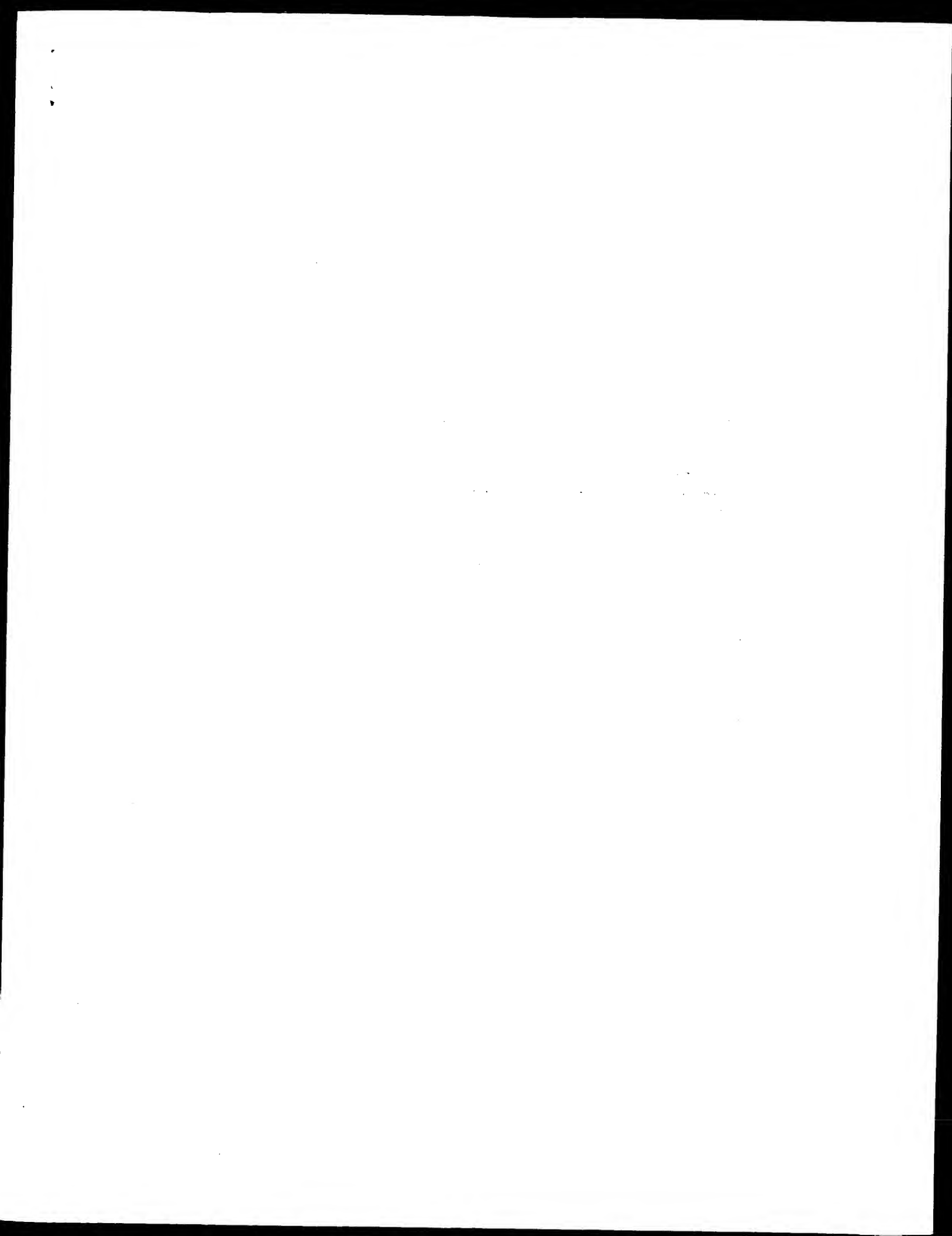
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX  
 SO Sequence 579 AA;

Query Match 100.0%; Score 151; DB 22; Length 579;  
 Best Local Similarly 100.0%; Pred. No. 3.3e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVVDFDDKYIVSASGDRITKW 29  
 ||||||||||||||||||  
 Db 397 VGHRAAVNVVDFDDKYIVSASGDRITKW 425

Search completed: April 11, 2003, 11:48:16  
 Job time : 9.93666 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 3.00244 Seconds  
(Without alignments)  
284.191 Million cell updates/sec

Title: US-09-601-168B-2\_COPY\_387\_415  
Perfect score: 151  
Sequence: 1 VGHRAAVNVDPDDKIVSASGDRITKW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	151	100.0	517	1 US-08-190-802A-30	Sequence 30, Appl
2	151	100.0	517	2 US-08-477-346-30	Sequence 30, Appl
3	151	100.0	517	4 US-08-473-089-30	Sequence 30, Appl
4	151	100.0	517	4 US-08-487-072A-30	Sequence 86, Appl
5	147	97.4	29	1 US-08-190-802A-86	Sequence 86, Appl
6	147	97.4	29	4 US-08-477-346-86	Sequence 86, Appl
7	147	97.4	29	4 US-08-473-089-86	Sequence 86, Appl
8	147	97.4	29	4 US-08-487-072A-86	Sequence 86, Appl
9	82	54.3	1146	4 US-08-914-999-6	Sequence 6, Appl
10	80	53.0	732	4 US-08-914-999-8	Sequence 8, Appl
11	76	50.3	514	1 US-08-190-802A-66	Sequence 66, Appl
12	76	50.3	514	4 US-08-477-346-66	Sequence 66, Appl
13	76	50.3	514	4 US-08-473-089-66	Sequence 66, Appl
14	76	50.3	514	4 US-08-487-072A-66	Sequence 66, Appl
15	75	49.7	209	3 US-08-899-578-6	Sequence 2, Appl
16	75	49.7	209	3 US-08-899-578-2	Sequence 2, Appl
17	74	49.0	33	1 US-08-190-802A-92	Sequence 92, Appl
18	74	49.0	33	4 US-08-473-089-92	Sequence 92, Appl
19	74	49.0	33	4 US-08-487-072A-92	Sequence 92, Appl
20	74	49.0	33	4 US-08-487-072A-10	Sequence 10, Appl
21	74	49.0	253	4 US-09-724-884-10	Sequence 2, Appl
22	74	49.0	253	4 US-09-724-884-2	Sequence 2, Appl
23	74	49.0	690	4 US-09-291-170A-2	Sequence 31, Appl
24	74	49.0	690	4 US-09-724-884-2	Sequence 31, Appl
25	74	49.0	906	1 US-08-190-802A-31	Sequence 31, Appl
26	74	49.0	906	4 US-08-477-346-31	Sequence 31, Appl
27	74	49.0	906	4 US-08-473-089-31	Sequence 31, Appl

28	74	49.0	906	4	US-08-487-072A-31	Sequence 31, Appl
29	72	47.7	640	4	US-09-177-165A-30	Sequence 30, Appl
30	72	47.7	2627	2	US-08-751-189-3	Sequence 3, Appl
31	72	47.7	2627	2	US-09-060-836-3	Sequence 3, Appl
32	72	47.7	2627	2	US-09-184-445-3	Sequence 4, Appl
33	72	47.7	2629	2	US-08-751-189-4	Sequence 4, Appl
34	71	47.0	2629	2	US-09-060-836-4	Sequence 4, Appl
35	71	47.0	2629	2	US-09-184-445-4	Sequence 4, Appl
36	70	46.4	798	1	US-08-190-802A-64	Sequence 64, Appl
37	70	46.4	798	1	US-08-190-802A-68	Sequence 64, Appl
38	70	46.4	798	2	US-08-308-818-2	Sequence 64, Appl
39	70	46.4	798	4	US-08-477-346-64	Sequence 64, Appl
40	70	46.4	798	4	US-08-473-089-64	Sequence 64, Appl
41	70	46.4	798	4	US-08-473-089-68	Sequence 64, Appl
42	70	46.4	798	4	US-08-487-072A-64	Sequence 64, Appl
43	70	46.4	798	4	US-08-487-072A-68	Sequence 64, Appl
44	69	45.7	29	1	US-08-190-802A-88	Sequence 88, Appl

## ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
Sequence 30, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Delinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R. 33, 875  
REGISTRATION NUMBER: 8600-0139  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0960  
TELEFAX: (415) 324-0980  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: Peptide  
HYPOTHETICAL: NO  
AMTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-190-802A-30

Query Match 100.0%; Score 151; DB 1; Length 517;  
Best local Similarity 100.0%; Pred. No. 3.3e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 VGHRAAVNVDPDDKIVSASGDRITKW 29

Db 351 VGHRAAVNVDFDDKXIVSASGDRITKW 379

## RESULT 2

US-08-473-346-30  
Sequence 30, Application US/0847346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theeotf  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-477-346-30

Query Match 100.0%; Score 151; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 3.5e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVDFDDKXIVSASGDRITKW 29  
Db 351 VGHRAAVNVDFDDKXIVSASGDRITKW 379

## RESULT 3

US-08-473-089-30  
Sequence 30, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theeotf  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-473-089-30

Query Match 100.0%; Score 151; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 3.5e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVDFDDKXIVSASGDRITKW 29  
Db 351 VGHRAAVNVDFDDKXIVSASGDRITKW 379

## RESULT 4

US-08-487-072A-30  
Sequence 30, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theeotf  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-487-072A-30

Query Match 100.0%; Score 151; DB 4;  
Best Local Similarity 100.0%; Pred. No. 3,5e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDFDDKTYIVSASGDRITIKW 29  
Db 351 VGHRAAVNVDFDDKTYIVSASGDRITIKW 379

RESULT 5  
US-08-190-802A-86  
Sequence 86, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0960  
TELEFAX: (415) 324-0860  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rIV, Fig. 13  
US-08-190-802A-86

Query Match 97.4%; Score 147; DB 1; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4.6e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHRAAVNVDFDDKTYIVSASGDRITIKW 29

Db 1 GHRAAVNVDFDDKTYIVSASGDRITIKW 28  
|||||

RESULT 6  
US-08-477-346-86  
Sequence 86, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rIV, Fig. 13  
US-08-477-346-86

Query Match 97.4%; Score 147; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4.6e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHRAAVNVDFDDKTYIVSASGDRITIKW 29  
Db 1 GHRAAVNVDFDDKTYIVSASGDRITIKW 28

RESULT 7  
US-08-473-089-86  
Sequence 86, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
FAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rIV, Fig. 13  
US-08-473-089-86

Query Match 97.4%; Score 147; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4,6e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GHRAAVNVVDFDDKTYIVASGDRITKYW 29  
|||||  
Db 1 GHRAAVNVVDFDDKTYIVASGDRITKYW 28

RESULT 8  
US-08-487-072A-86  
Sequence 86, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
FAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rIV, Fig. 13  
US-08-487-072A-86

Query Match 97.4%; Score 147; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4,6e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GHRAAVNVVDFDDKTYIVASGDRITKYW 29  
|||||  
Db 1 GHRAAVNVVDFDDKTYIVASGDRITKYW 28

RESULT 9  
US-08-914-999-6  
Sequence 6, Application US/08914999  
Patent No. 6346406  
GENERAL INFORMATION:  
APPLICANT: Ryazanov, Alexey G.  
APPLICANT: Halt, William N.  
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)  
TITLE OF INVENTION: AND METHODS OF USE THEREFOR  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
FLOOR  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/914,999  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 601-1-078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
FAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1146 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Dictyostellium discoideum  
US-08-914-999-6

Query Match 54.3%; Score 82; DB 4; Length 1146;  
Best Local Similarity 55.2%; Pred. No. 0.00035;  
Matches 16; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 VGHRAAVNVVDPPDKYIVSASGDRITIKW 29  
DB 951 VGHTEVNCVANEKYLFGSCSYDKITIKW 979

RESULT 10

US-08-914-999-8  
Sequence 8, Application US/08914999  
Patent No. 6346406  
GENERAL INFORMATION:  
APPLICANT: Ryazanov, Alexey G.  
APPLICANT: Hailt, William N.  
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)  
TITLE OF INVENTION: AND METHODS OF USE THEREFOR  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
STREET: floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/914,999  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION/DOCKET NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 601-1-078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Dictyostellium discoideum  
US-08-914-999-8  
Query Match 53.0%; Score 80; DB 4; Length 732;  
Best Local Similarity 53.6%; Pred. No. 0.00042;  
Matches 15; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 2 GHRRAVNVVDPPDKYIVSASGDRITIKW 29  
DB 540 GHRPVTHTVLNDKYLFGSGSDKITIKW 567

RESULT 11

US-08-190-802A-66  
Sequence 66, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Delinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION/DOCKET NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49

US-08-190-802A-66  
Query Match 50.3%; Score 76; DB 1; Length 514;  
Best Local Similarity 56.7%; Pred. No. 0.0011;  
Matches 17; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

OY 2 GHRRAVNVVDPPDKYIVSASGDRITIKW 29  
DB 400 GHRKLVNVHVFSPDGRYIVSASFDSIKIW 429

RESULT 12

US-08-477-346-66  
Sequence 66, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49  
US-08-473-089-66

Query Match 50.3%; Score 76; DB 4; Length 514;  
Best Local Similarity 56.7%; Pred. No. 0.0011;  
Matches 17; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

OY 2 GHRAAVWVDF--DDKYIVASGDRITKYW 29  
||: ||| | | : ||||| | : ||: |  
Db 400 GHOKLVNHVAFSPDGRIVASAFDINSIKLM 429

RESULT 13  
US-08-473-089-66  
Sequence 66, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49  
US-08-473-089-66

Query Match 50.3%; Score 76; DB 4; Length 514;

Best Local Similarity 56.7%; Pred. No. 0.0011;  
Matches 17; Conservative 4; Mismatches 7; Indels 2; Gaps 1;  
OY 2 GHRAAVWVDF--DDKYIVASGDRITKYW 29  
||: ||| | | : ||||| | : ||: |  
Db 400 GHOKLVNHVAFSPDGRIVASAFDINSIKLM 429

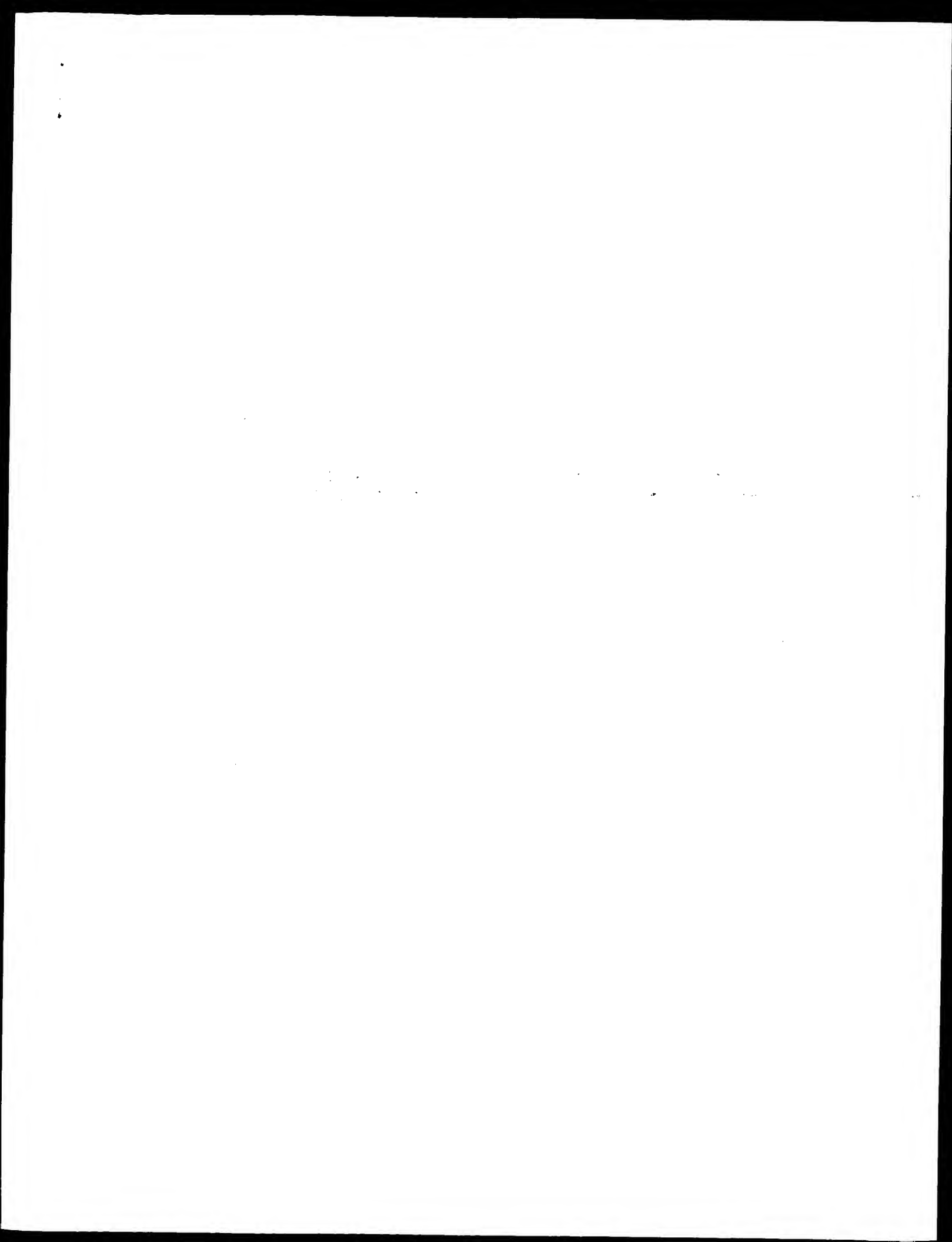
RESULT 14  
US-08-487-072A-66  
Sequence 66, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49  
US-08-487-072A-66

Query Match 50.3%; Score 76; DB 4; Length 514;  
Best Local Similarity 56.7%; Pred. No. 0.0011;  
Matches 17; Conservative 4; Mismatches 7; Indels 2; Gaps 1;  
OY 2 GHRAAVWVDF--DDKYIVASGDRITKYW 29  
||: ||| | | : ||||| | : ||: |  
Db 400 GHOKLVNHVAFSPDGRIVASAFDINSIKLM 429

RESULT 15  
US-08-899-578-6  
Sequence 6, Application US/08899578  
Patent No. 6087153  
GENERAL INFORMATION:  
APPLICANT: Greenwald, Iva  
APPLICANT: Hubbard, E. Jane  
TITLE OF INVENTION: SEL-10 AND USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP







Fri Apr 11 13:31:11 2003

us-09-601-168b-2-copy\_387\_415.rapb

Page 1

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:49:06 ; Search time 2.96711 Seconds  
(without alignments)  
597.532 Million cell updates/sec

Title: US-09-601-168b-2-copy\_387\_415  
Perfect score: 151  
Sequence: 1 VGHRAAVNVDFDXYIVSASGDRITKW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	151	100.0	569	US-10-038-010-8	Sequence 8, Appl1
2	151	100.0	569	US-10-042-417-2	Sequence 2, Appl1
3	83	55.0	678	US-09-801-368-314	Sequence 314, App
4	82	54.3	1146	US-09-832-292-10	Sequence 10, Appl
5	82	54.3	1146	US-09-994-485-6	Sequence 6, Appl1
6	80	53.0	732	US-09-994-485-8	Sequence 8, Appl1
7	77	51.0	296	US-10-083-357-1300	Sequence 1300, Ap
8	76	50.3	732	US-09-832-292-12	Sequence 12, Appl
9	74	49.0	744	US-09-925-300-1347	Sequence 1347, Ap
10	72	47.7	485	US-10-132-744A-6	Sequence 6, Appl1
11	72	47.7	640	US-10-060-019-30	Sequence 30, Appl
12	70	46.4	422	US-10-042-417-4	Sequence 4, Appl1
13	69	45.7	159	US-09-774-639-247	Sequence 247, App
14	69	45.7	159	US-09-969-730-345	Sequence 345, App
15	69	45.7	261	US-10-132-744A-4	Sequence 4, Appl1
16	69	45.7	484	US-10-132-744A-2	Sequence 2, Appl1
17	69	45.7	540	US-09-213-888-7	Sequence 7, Appl1
18	69	45.7	540	US-09-213-888-10	Sequence 10, Appl
19	69	45.7	540	US-09-328-877A-7	Sequence 7, Appl1

20	69	45.7	540	US-09-328-877A-10	Sequence 10, Appl
21	69	45.7	545	US-09-213-888-6	Sequence 6, Appl1
22	69	45.7	545	US-09-328-877A-6	Sequence 6, Appl1
23	69	45.7	553	US-09-213-888-5	Sequence 5, Appl1
24	69	45.7	553	US-09-328-877A-5	Sequence 5, Appl1
25	69	45.7	559	US-09-213-888-9	Sequence 9, Appl1
26	69	45.7	559	US-09-328-877A-9	Sequence 8, Appl1
27	69	45.7	589	US-09-213-888-8	Sequence 8, Appl1
28	69	45.7	589	US-09-328-877A-8	Sequence 4, Appl1
29	69	45.7	592	US-09-213-888-4	Sequence 4, Appl1
30	69	45.7	592	US-09-328-877A-4	Sequence 21, Appl
31	69	45.7	626	US-09-213-888-21	Sequence 21, Appl
32	69	45.7	626	US-09-328-877A-21	Sequence 3, Appl1
33	69	45.7	627	US-09-213-888-3	Sequence 3, Appl1
34	69	45.7	627	US-09-328-877A-3	Sequence 27, Appl
35	69	45.7	666	US-09-213-888-27	Sequence 27, Appl
36	69	45.7	666	US-09-328-877A-27	Sequence 25, Appl
37	69	45.7	669	US-09-213-888-25	Sequence 25, Appl
38	69	45.7	669	US-09-328-877A-25	Sequence 449, App
39	69	45.0	521	US-09-764-853-449	Sequence 29, Appl
40	68	45.0	779	US-10-060-019-29	Sequence 10, Appl
41	67.5	44.7	372	US-09-967-552A-10	Sequence 10, Appl
42	67	44.4	1356	US-10-077-111-10	Sequence 1445, Ap
43	66	43.7	114	US-09-796-692-1445	Sequence 1541, Ap
44	66	43.7	114	US-09-796-692-1541	Sequence 1895, Ap
45	66	43.7	114	US-09-796-692-1895	

#### ALIGNMENTS

```
RESULT 1
US-10-038-010-8
Sequence 8, Application US/10038010
Publication No. US20030040089A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
APPLICANT: Pierre, Legrain
TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
FILE REFERENCE: B4767A
CURRENT APPLICATION NUMBER: US/10/038, 010
PRIORITY FILING DATE: 2002-07-23
PRIOR APPLICATION NUMBER: US 60/259, 377
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 569
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Beta-TGPI
LOCATION: (1)..(569)
OTHER INFORMATION:
US-10-038-010-8
Query Match 100.0% Score 151; DB 9; Length 569;
Best Local Similarity 100.0% Pred. No. 2.9e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 VGHRAAVNVDFDXYIVSASGDRITKW 29
Db 387 VGHRAAVNVDFDXYIVSASGDRITKW 415
RESULT 2
US-10-042-417-2
Sequence 2, Application US/10042417
Patent No. US20020123082A1
GENERAL INFORMATION:
APPLICANT: Pasano, M.
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
```

FILE REFERENCE: 5914-090-999  
CURRENT APPLICATION NUMBER: US/10/042,417  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-042-417-2

Query Match 100.0%; Score 151; DB 12; Length 569;  
Best Local Similarity 100.0%; Pred. No. 2,9e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDPDDKYIVSASGDRITKW 29  
Db 387 VGHRAAVNVDPDDKYIVSASGDRITKW 415

RESULT 3  
US-09-801-368-314  
Sequence 314, Application US/09801368  
Patent No. US20020128250A1  
GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Call, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 314  
LENGTH: 678  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
US-09-801-368-314

Query Match 55.0%; Score 83; DB 10; Length 678;  
Best Local Similarity 50.0%; Pred. No. 0.00011;  
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 GHRRAAVNVDPDDKYIVSASGDRITKW 29  
Db 387 GHRRAAVNVDPDDKYIVSASGDRITKW 414

RESULT 4  
US-09-832-292-10  
Sequence 10, Application US/09832292  
Patent No. US20020177205A1  
GENERAL INFORMATION:  
APPLICANT: Ryzanov, Alexey  
TITLE OF INVENTION: MAMMALIAN ALPHA-KINASE PROTEINS, NUCLEIC ACIDS AND  
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USES THEREOF

FILE REFERENCE: 601-1-098CIP  
CURRENT APPLICATION NUMBER: US/09/832,292  
CURRENT FILING DATE: 2001-04-10  
PRIOR APPLICATION NUMBER: 09/632,131  
PRIOR FILING DATE: 2001-08-03  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 10  
LENGTH: 1146  
TYPE: PRT  
ORGANISM: Dictyostelium discoideum  
US-09-832-292-10

Query Match 54.3%; Score 82; DB 9; Length 1146;  
Best Local Similarity 55.2%; Pred. No. 0.00029;  
Matches 16; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDPDDKYIVSASGDRITKW 29  
Db 951 VGHTEVNCVANKETLFCSTYDKITKW 979

RESULT 5  
US-09-994-485-6  
Sequence 6, Application US/09994485  
Patent No. US20020142429A1  
GENERAL INFORMATION:  
APPLICANT: Ryzanov, Alexey G.  
Halt, William N.  
Pavur, Karen S.  
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)  
AND METHODS OF USE THEREFOR  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSER: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/994,485  
FILING DATE: 27-Nov-02 US20020142429A1-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/914,999  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq, David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 601-1-078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1146 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Dictyostelium discoideum  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-994-485-6

Query Match 54.3%; Score 82; DB 10; Length 1146;  
Best Local Similarity 55.2%; Pred. No. 0.00029;  
Matches 16; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 VGHRAAVNVDFDDKYIVSASGDRITKW 29  
||| | : : : : :  
Db 951 VGHTEVNCVANEKYLFCSDYDKITKW 979

RESULT 6  
US-09-994-485-8  
Sequence 8, Application US/09994485  
Patent No. US20020142429A1

GENERAL INFORMATION:  
APPLICANT: Ryazanov, Alexey G.  
Halt, William N.  
Pavur, Karen S.

TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)  
AND METHODS OF USE THEREFOR

NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/994,485  
FILING DATE: 27-NO. US20020142429A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/914,999

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 601-1-078

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5600

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 732 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Dictyostellium discoideum

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Query Match 53.0%; Score 80; DB 10; Length 732;  
Best Local Similarity 53.6%; Pred. No. 0.00035;  
Matches 15; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 GHRAAVNVDFDDKYIVSASGDRITKW 29  
||| | : : : : :  
Db 540 GHDKPVTHVLNDKYLFGSSSDKITKW 567

RESULT 7  
US-10-083-357-1300  
Sequence 1300, Application US/10083357  
Publication No. US20030054370A1

GENERAL INFORMATION:  
APPLICANT: Qiangdong zeng et al.  
TITLE OF INVENTION: Systemic Discovery of New Genes

FILE REFERENCE: 032796-090

CURRENT APPLICATION NUMBER: US/10/083,357

CURRENT FILING DATE: 2002-02-27

NUMBER OF SEQ ID NOS: 1346

SEQ ID NO 1300

LENGTH: 296

TYPE: PRT

ORGANISM: Saccharomyces cerevisiae

US-10-083-357-1300

Query Match 51.0%; Score 77; DB 9; Length 296;  
Best Local Similarity 50.0%; Pred. No. 0.00035;  
Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 2 GHRAAVNVDFDDKYIVSASGDRITKW 29  
||| | : : : : :  
Db 182 GHRALVGLGLSDKFLVSASVDGSIKWC 209

RESULT 8  
US-09-832-292-12  
Sequence 12, Application US/09832292  
Patent No. US20020177205A1

GENERAL INFORMATION:  
APPLICANT: Ryazanov, Alexey

TITLE OF INVENTION: MAMMALIAN ALPHA-KINASE PROTEINS, NUCLEIC ACIDS AND

FILE REFERENCE: 601-1-098CIP

CURRENT APPLICATION NUMBER: US/09/832,292

CURRENT FILING DATE: 2001-04-10

PRIOR APPLICATION NUMBER: 09/632,131

PRIOR FILING DATE: 2001-08-03

NUMBER OF SEQ ID NOS: 45

SOFTWARE: Patentin version 3.1

SEQ ID NO 12

LENGTH: 732

TYPE: PRT

ORGANISM: Dictyostellium discoideum

US-09-832-292-12

Query Match 50.3%; Score 76; DB 9; Length 732;  
Best Local Similarity 53.6%; Pred. No. 0.0014;  
Matches 15; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 2 GHRAAVNVDFDDKYIVSASGDRITKW 29  
||| | : : : : :  
Db 540 GHDKPVTHVLNDKYLFGSSSDKITKW 567

RESULT 9  
US-09-925-300-1347  
Sequence 1347, Application US/09925300  
Patent No. US20020151681A1

GENERAL INFORMATION:  
APPLICANT: Craig Rosen,

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05988

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

NUMBER OF SEQ ID NOS: 1890

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 1347

LENGTH: 744

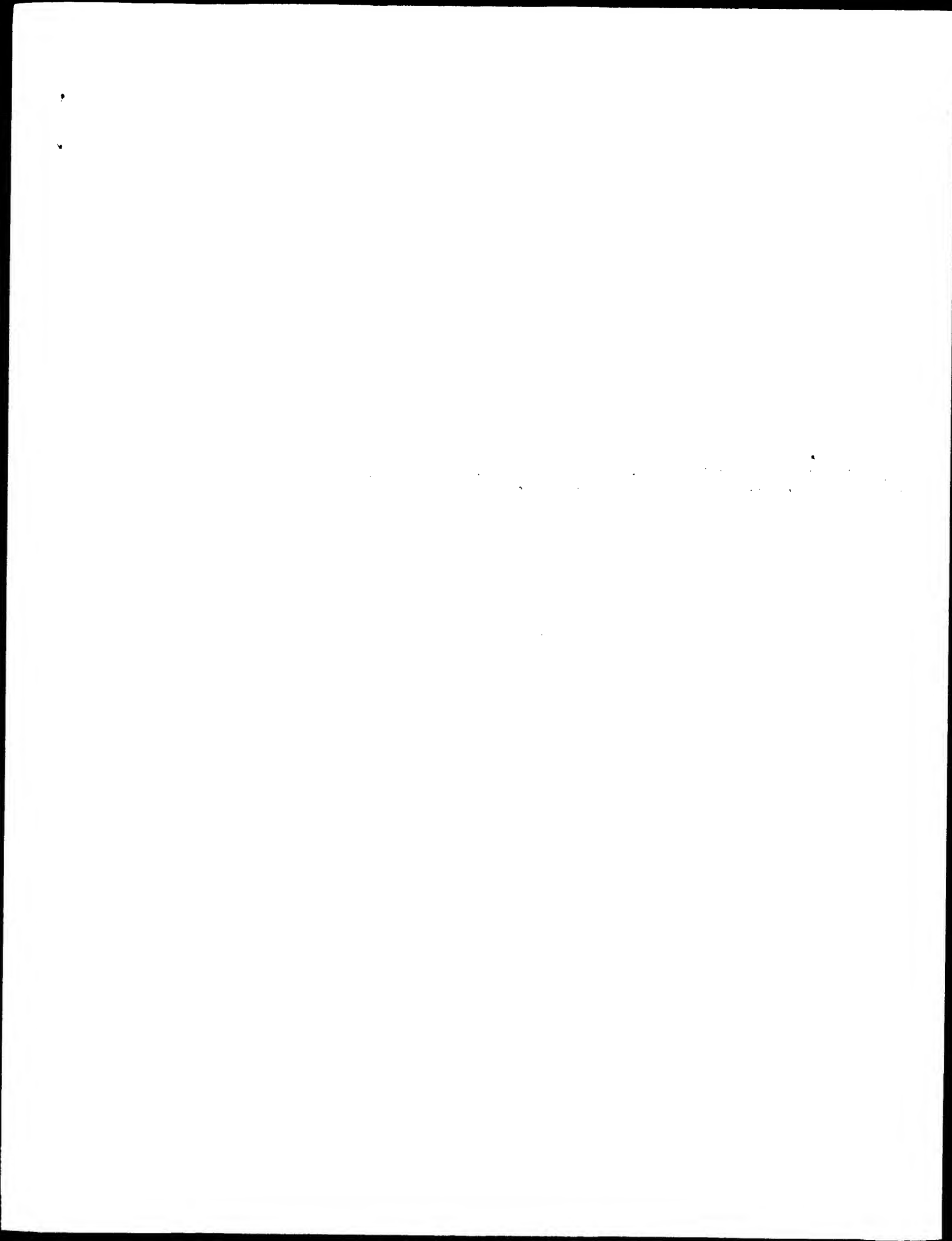
TYPE: PRT

ORGANISM: Homo sapiens

100

100







GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 3.35566 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_387\_415  
Sequence score: 151  
1 VGHRAAVNVVDFDDKTIYVSASGDRITIKVW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR-73:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	100.0	518	2 B48088	beta-transducin re
2	148	98.0	701	2 T16607	hypothetical prote
3	85	56.3	650	2 T46660	sulfur controller
4	83	55.0	934	2 AG1889	WD-40 repeat prote
5	82	54.3	1146	2 A55532	myosin-heavy chain
6	81	53.6	925	2 T29585	hypothetical prote
7	79	52.3	1551	2 AB2410	WD-repeat protein
8	78	51.7	1747	2 AC1842	WD-40 repeat prote
9	77	51.0	185	2 S50569	hypothetical prote
10	77	51.0	319	2 S54578	hypothetical prote
11	76	50.3	515	2 S19487	hypothetical prote
12	75	49.7	530	2 T20360	hypothetical prote
13	75	49.7	571	2 T20359	hypothetical prote
14	75	48.7	579	2 T22703	hypothetical prote
15	75	48.7	1258	2 A12155	WD-repeat protein
16	74	49.0	535	2 S48268	probable membrane
17	74	49.0	906	2 S35342	Golgi-associated p
18	74	49.0	906	2 S35312	coatomer complex b
19	74	49.0	1189	2 AH2154	WD-repeat protein
20	74	49.0	1711	2 AD1842	WD-40 repeat prote
21	73	48.3	267	2 S62507	hypothetical trp-a
22	73	48.3	325	2 T06784	GTP-binding protei
23	73	48.3	325	2 T09613	probable GTP-bind
24	73	48.3	327	2 S48839	guanine nucleotide
25	73	48.3	327	2 T50211	WD-repeat protein
26	72	47.7	334	2 T03764	protein RMD - rice
27	72	47.7	640	2 S49932	MTJ30 protein - ye
28	72	47.7	876	2 T51507	WD40-repeat protei
29	71	47.0	303	2 S45461	hypothetical prote

30	71	47.0	473	2 T33805	hypothetical prote
31	71	47.0	943	2 S59317	DIP2 protein - yea
32	71	47.0	1683	2 AF2071	WD-40 repeat prote
33	71	47.0	1693	2 S76086	beta transducin-li
34	71	47.0	2629	2 T30987	telomerase-associa
35	71	47.0	2629	2 T32735	telomerase-associa
36	70	46.4	290	2 T02300	GTP-binding regula
37	70	46.4	316	2 S57839	CPC2 protein - Neu
38	70	46.4	317	2 T46032	WD-40 repeat regul
39	70	46.4	372	2 T15181	hypothetical prote
40	70	46.4	559	2 AB2202	hypothetical prote
41	70	46.4	586	2 T38992	WD-40 repeat regul
42	70	46.4	798	2 S34023	TATA box-binding p
43	70	46.4	889	2 B55123	coatomer complex b
44	70	46.4	1269	2 T00443	hypothetical prote
45	70	46.4	1389	2 H84914	probable WD-40 rep

#### ALIGNMENTS

##### RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-Trop

C:Species: Xenopus laevis (African clawed frog)

C>Date: 26-May-1994 #sequence\_rev150 26-May-1994 #text\_change 21-Jul-2000

C:Accession: B48088

R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: A48088; M01D:93330289; PMID:8393141

A:Accession: B48088

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SP>

A:Cross-references: GB:098268; NID:9295542; PIDN:AAA02810.1; PID:9295543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <WD1>

Query Match	100.0%	Score 151	DB 2	Length 518
Best Local Similarity	100.0%	Pred. No. 2.8e-14		
Matches 29	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Db	352 VGHRAAVNVVDFDDKTIYVSASGDRITIKVW 380			

##### RESULT 2

T16607 hypothetical protein K10B2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_rev150 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16607

R:Miller, N.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid K10B2.

A:Reference number: Z18545

A:Accession: T16607

A:Status: preliminary; translated from GH/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-701 <MIL>

A:Cross-references: EMBL:U28730; NID:9860694; PID:9860695; PIDN:AAA68258.1; CESP:K10B

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K10B2.1

A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match	98.0%	Score 148	DB 2	Length 701
Best Local Similarity	96.6%	Pred. No. 1.1e-13		
Matches 28	Conservative 1	Mismatches 0	Indels 0	Gaps 0

OY 1 VGHRAAVNVDFDDKYIVSASGDRITKW 29  
 DB 378 VGHRAAVNVDFDDKYIVSASGDRITKW 406

### RESULT 3

146660  
 sulfur controller-2 protein [imported] - Neurospora crassa  
 C:Species: Neurospora crassa  
 C>Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
 C:Accession: T46660  
 R:Kumar, A.; Paletta, J.V.  
 Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995  
 A:Title: The sulfur controller-2 negative regulatory gene of Neurospora crassa encodes a  
 A:Reference number: 223121; MUID:95241499; PMID:7724564  
 A:Accession: T46660  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-650 <KUM>  
 A:Cross-references: EMBL:U17251; NID:g806757; PIDN:AAA68968.1; PID:g806758  
 C:Genetics:  
 A:Gene: scon-2  
 A:Map position: 3  
 A:Introns: 75/3; 319/1; 354/1  
 C:Function:  
 A:Description: negatively regulates sulfur structural gene expression  
 A:Note: scon-2+ expression is dependent on CYS3 function and the binding of CYS3 to the  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 56.3%; Score 85; DB 2; Length 650;  
 Best Local Similarity 51.7%; Pred. No. 0.0002;  
 Matches 15; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDFDDKYIVSASGDRITKW 29  
 DB 331 VGHTRAGIRALQFDDSKLISGSDHTIKW 359

### RESULT 4

AGI889  
 WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AGI889  
 R:kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriugu  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AGI889  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-934 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA872622.1; PID:917130010; GSPDB:GN00179  
 C:Genetics:  
 A:Experimental source: strain PCC 7120  
 A:Gene: al10664

Query Match 55.0%; Score 83; DB 2; Length 934;  
 Best Local Similarity 56.7%; Pred. No. 0.00059;  
 Matches 17; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

OY 2 GHRAAVNVDFDDKYIVSASGDRITKW 29  
 DB 530 GHQNGVSVTFSPDGKLIATASGDRITKW 559

### RESULT 5

A55532  
 myosin-heavy-chain kinase (EC 2.7.1.129) A - slime mold (Dictyostelium discoideum)  
 C:Species: Dictyostelium discoideum

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-May-2000  
 C:Accession: A55532  
 R:Futley, L.M.; Medley, Q.G.; Cole, G.P.; Egelhoff, T.T.  
 J. Biol. Chem. 270, 523-529, 1995  
 A:Title: Structural analysis of myosin heavy chain kinase A from Dictyostelium. Evid  
 the beta-subunit of heterotrimeric G proteins.  
 A:Reference number: A55532; MUID:95122486; PMID:7822274

A:Accession: A55532  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1146 <FUT>  
 A:Cross-references: GB:U16856; NID:g608519; PIDN:AAA66070.1; PID:g608520  
 C:Genetics:  
 A:Introns: #status absent  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 C:Keywords: autophosphorylation; coiled coil; multimer; phosphoprotein; phosphotransf  
 F:865-898/Domain: WD repeat homology <WD1>  
 F:950-981/Domain: WD repeat homology <WD2>  
 F:991-1022/Domain: WD repeat homology <WD3>  
 F:1031-1062/Domain: WD repeat homology <WD4>  
 F:1071-1102/Domain: WD repeat homology <WD5>  
 F:1112-1143/Domain: WD repeat homology <WD6>

Query Match 54.3%; Score 82; DB 2; Length 1146;  
 Best Local Similarity 55.2%; Pred. No. 0.001;  
 Matches 16; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDFDDKYIVSASGDRITKW 29  
 DB 951 VGHTEVNCVAVNEKYLEFSCYDRTIKW 979

### RESULT 6

T29585  
 hypothetical protein F55F8.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 26-May-2000  
 C:Accession: T29585  
 R:Gattung, S.; Scheet, P.; Kemp, K.  
 submitted to the EMBL data library, November 1996  
 A:Description: The sequence of C. elegans cosmid F55F8.  
 A:Reference number: Z20647  
 A:Accession: T29585  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-925 <GAT>  
 A:Cross-references: EMBL:U80447; PIDN:AA837807.1; GSPDB:GN00019; CESP:F55F8.3  
 C:Genetics:  
 A:Experimental source: strain Bristol N2; clone F55F8  
 A:Gene: CESP:F55F8.3  
 A:Map position: 1  
 A:Introns: 6/3; 44/2; 76/1; 109/2; 159/2; 204/3; 327/1; 469/3; 617/1; 775/1; 825/3; 8  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 53.6%; Score 81; DB 2; Length 925;  
 Best Local Similarity 50.0%; Pred. No. 0.0011;  
 Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRAAVNVDFDDKYIVSASGDRITKW 29  
 DB 486 GHESAISIDHGNHIVSGSWDRITKW 513

### RESULT 7

AB2410  
 WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AB2410  
 R:kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriugu  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata  
 DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AB2410  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1551 <KOR>  
A:Cross-references: GB:BA000019; PIDN:BA076533.1; PID:g17133971; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all4834

Query Match 52.3%; Score 79; DB 2; Length 1551;  
Best Local Similarity 46.7%; Pred. No. 0.004;  
Matches 14; Conservative 7; Mismatches 7; Indels 2; Gaps 1;

OY 2 GHRAAVNVDF--DDKYIYASGDRITIKW 29  
DB 1484 GHRGAVRSISIPDQYIATASDDRTVKLM 1513

## RESULT 8

WD:40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp.  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AC1842  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuriiz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AC1842  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1747 <KOR>  
A:Cross-references: GB:BA000019; PIDN:BA077807.1; PID:g17135261; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all0283

Query Match 51.7%; Score 78; DB 2; Length 1747;  
Best Local Similarity 53.3%; Pred. No. 0.0064;  
Matches 16; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

OY 2 GHRAAVNVDF--DDKYIYASGDRITIKW 29  
DB 1135 GHRGAVNVASFDPGKVLASASDDRTVKLM 1164

## RESULT 9

hypothetical protein YER066w - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 19-Apr-2002  
C:Accession: S50569  
R:Dietrich, F.S.

submitted to the EMBL Data Library, December 1994  
A:Description: The sequence of S. cerevisiae lambda clones 6592, 4678, 4742, and 3612.  
A:Reference number: S50427  
A:Accession: S50569  
A:Molecule type: DNA  
A:Residues: 1-185 <DIE>  
A:Cross-references: EMBL:U01813; NID:g1381127; PIDN:AAB64602.1; PID:g603302; MIPS:YER066  
C:Genetics:  
A:Cross-references: SGD:S0000868  
A:Map position: 5R  
C:Superfamily: Saccharomyces hypothetical protein YER066w; WD repeat homology  
C:Keywords: duplication  
F:7-40/Domain: WD repeat homology <WD1>  
F:69-100/Domain: WD repeat homology <WD2>

Query Match 51.0%; Score 77; DB 2; Length 185;

Best Local Similarity 50.0%; Pred. No. 0.00077;  
Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRAAVNVDFDDKYIYASGDRITIKW 29  
DB 71 GHRALVGLISLSDKFLVASVDSISRCW 98

## RESULT 10

hypothetical protein YMR16c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YMR16c, 14b; hypothetical protein YMR16c, 15c  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 26-May-2000  
C:Accession: S54578; S54577  
R:Hunt, S.; Bowman, S.

submitted to the EMBL Data Library, May 1995  
A:Reference number: S54510  
A:Accession: S54578  
A:Molecule type: DNA  
A:Residues: 1-319 <HUN>  
A:Cross-references: EMBL:Z49702; NID:g817859; PIDN:CAA89754.1; PID:g817876; MIPS:YMR1  
A:Experimental source: strain AB972  
A:Accession: S54577  
A:Molecule type: DNA  
A:Residues: 221-319 <HUN>  
A:Cross-references: EMBL:Z49702; NID:g817859; PIDN:CAA89753.1; PID:g817875; MIPS:YMR1  
A:Experimental source: strain AB972  
C:Genetics:  
A:Gene: SGD:BEL1  
A:Cross-references: SGD:S0004722; MIPS:YMR16c  
A:Map position: 13R  
A:Introns: 179/3  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
F:13-47/Domain: WD repeat homology <WD1>  
F:61-94/Domain: WD repeat homology <WD2>  
F:103-136/Domain: WD repeat homology <WD3>  
F:144-183/Domain: WD repeat homology <WD4>  
F:192-225/Domain: WD repeat homology <WD5>  
F:235-265/Domain: WD repeat homology <WD6>  
F:283-315/Domain: WD repeat homology <WD7>

Query Match 51.0%; Score 77; DB 2; Length 319;  
Best Local Similarity 54.8%; Pred. No. 0.0014;  
Matches 17; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

OY 1 VGHRAVNVDFDDK--YIVSASGDRITIKW 29  
DB 104 VGHKSDVMSVDIDKAKSMIISGSRDKITIKW 134

## RESULT 11

hypothetical protein YCR072c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 31-Mar-1992 #sequence\_revision 04-Dec-1992 #text\_change 19-Apr-2002  
C:Accession: S19487; S26557  
R:Ballester, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.

submitted to the Protein Sequence Database, March 1992  
A:Reference number: S19486  
A:Accession: S19487  
A:Molecule type: DNA  
A:Residues: 1-484, 1', 486, 7'KL', 490, 492-515 <BAL1>  
A:Cross-references: EMBL:X59720; MIPS:YCR072c  
A:Note: this sequence has been revised in reference S26557  
A:Reference number: S26587  
A:Accession: S26557  
A:Molecule type: DNA  
A:Residues: 481-503 <BAL2>  
A:Cross-references: EMBL:X59720; MIPS:YCR072c  
C:Genetics:

hypothetical protein YCR072c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 31-Mar-1992 #sequence\_revision 04-Dec-1992 #text\_change 19-Apr-2002  
C:Accession: S19487; S26557  
R:Ballester, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.

submitted to the Protein Sequence Database, October 1992  
A:Reference number: S26587  
A:Accession: S26557  
A:Molecule type: DNA  
A:Residues: 481-503 <BAL2>  
A:Cross-references: EMBL:X59720; MIPS:YCR072c  
C:Genetics:

Query Match 51.0%; Score 77; DB 2; Length 319;  
Best Local Similarity 54.8%; Pred. No. 0.0014;  
Matches 17; Conservative 4; Mismatches 8; Indels 2; Gaps 1;



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: April 10, 2003, 12:51:26 ; Search time 2.08404 Seconds  
(without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_387\_415

Perfect score: 151  
Sequence: 1 VGHRAAVNVVDFDDKYIVASGDRTIKVW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	100.0	518	1 TRCB_XENLA	Q91854 xenopus lae
2	151	100.0	542	1 FM1B_HUMAN	Q9UKX1 homo sapien
3	151	100.0	605	1 FM1A_HUMAN	Q9Y297 homo sapien
4	148	98.0	665	1 LI23_CAEEL	Q09990 caenorhabdi
5	85	56.3	650	1 SCOB_NEUCR	Q01277 neurospora
6	83	55.0	678	1 SCOB_EMENT	Q00659 ementella
7	82	54.3	1146	1 KHALDICTI	P42527 dictyosteli
8	81	53.6	910	1 PMW2_CAEEL	P13341 caenorhabdi
9	80	53.0	732	1 KMBB_DICDI	P40042 dictyosteli
10	77	51.0	185	1 YEP6_YEAST	P38011 saccharomyc
11	77	51.0	319	1 GBLP_YEAST	P25382 saccharomyc
12	76	50.3	515	1 YCWL_YEAST	P90794 caenorhabdi
13	75	49.7	571	1 YRQ_CAEEL	Q93794 caenorhabdi
14	75	49.7	579	1 SETO_CAEEL	O62621 dictyosteli
15	75	49.7	914	1 COPP_DROME	O8YTC2 anabadena sp
16	75	49.7	1258	1 YS00_ANASP	P38262 saccharomyc
17	74	49.0	535	1 SIR2_YEAST	O35142 rattus norv
18	74	49.0	904	1 COPP_RAT	P35605 bos taurus
19	74	49.0	905	1 COPP_BOVIN	P35606 homo sapien
20	74	49.0	905	1 COPP_HUMAN	O55029 mus musculu
21	74	49.0	905	1 COPP_MOUSE	O24076 medicago sa
22	73	48.3	325	1 GBLP_MEDSA	Q34836 glycyline max
23	73	48.3	325	1 GBLP_SOYBN	Q34456 arabidopsi
24	73	48.3	327	1 GBLP_ARATH	Q33336 brassica na
25	73	48.3	327	1 GBLP_BRANA	O09855 schizosacch
26	73	48.3	506	1 PORB_SCHPO	P53699 candida alb
27	73	48.3	684	1 CC4_CANAL	P49027 oryza sativ
28	72	47.7	334	1 GBLP_ORYSA	P49027 oryza sativ
29	72	47.7	640	1 MT30_YEAST	P33914 saccharomyc
30	71	47.0	303	1 YNA6_YEAST	P43138 saccharomyc
31	71	47.0	943	1 DIP2_YEAST	Q12220 saccharomyc
32	71	47.0	1683	1 YL24_ANASP	O8YV57 anabadena sp
33	71	47.0	1693	1 Y163_SYNY3	Q55563 synchocyst

## ALIGNMENTS

RESULT 1	ID	TRCB_XENLA	STANDARD:	PRT:	518 AA.	
AC	Q91854	P70037	P70038:			Q01369 neurospora
DT	15-JUL-1998	(Rel. 36, Last sequence update)				Q9UKX1 homo sapien
DT	15-JUL-1998	(Rel. 36, Last sequence update)				Q9UG88 schizosacch
DT	15-JUN-2002	(Rel. 41, Last annotation update)				P38129 saccharomyc
DE	Beta-TrCP (beta-transducin repeat-containing protein).					P41811 saccharomyc
GN	FBXW1 OR BTRCP.					P36037 saccharomyc
OS	Xenopus laevis (African clawed frog).					Q92828 homo sapien
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					P87053 schizosacch
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;					O09715 schizosacch
OC	Xenopodinae; Xenopus.					Q9P713 schizosacch
OX	NCBI_TaxID=8335;					O14170 schizosacch
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=9330289; PubMed=8393141;					
RA	Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;					
RT	"Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in					
RT	anaphase are rescued by Xenopus CDNs encoding N-ras or a protein					
RT	with beta-transducin repeats."					
RL	Mol. Cell. Biol. 13:4953-4966(1993).					
RN	[2]					
RP	SEQUENCE OF 302-518 FROM N.A.					
RX	MEDLINE=97109804; PubMed=8952061;					
RA	Hudson J.W., Alarcon V.B., Ellinson R.P.;					
RT	"Identification of new localized RNAs in the Xenopus oocyte by					
RT	differential display PCR."					
RL	Dev. Genet. 19:190-198(1996).					
CC	-1- FUNCTION: Probably recognizes and binds to some phosphorylated					
CC	proteins and promotes their ubiquitination and degradation.					
CC	-1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX					
CC	(BY SIMILARITY).					
CC	-1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-					
CC	MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC					
CC	GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION, DO					
CC	NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR					
CC	TADPOLE EMBRYO.					
CC	-1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.					
CC	-1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).					
CC	-1- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>					
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).					
CC	-----					
DR	EMBL: M98268; AAA02810.1; -					
DR	EMBL: U63921; AAB49671.1; -					
DR	EMBL: U63922; AAB49672.1; -					
DR	InterPro: IPR001810; F-box.					
DR	InterPro: IPR001680; WD40.					
DR	Pfam: PF00400; WD40; 7.					

DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR Prodom: PD00018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS0181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR UBL conjugation pathway; Repeat: WD repeat.  
 KW UBL conjugation pathway; Repeat: WD repeat.  
 FT DOMAIN 119 157 F-BOX.  
 FT REPEAT 230 258 WD 1.  
 FT REPEAT 270 298 WD 2.  
 FT REPEAT 310 338 WD 3.  
 FT REPEAT 353 381 WD 4.  
 FT REPEAT 393 421 WD 5.  
 FT REPEAT 433 461 WD 6.  
 FT REPEAT 482 510 WD 7.  
 FT REPEAT 502 518 WD 7.  
 FT CONFLICT 302 304 GEM -> EFR (IN REF. 2).  
 FT CONFLICT 516 518 GLA -> AAH (IN REF. 2).  
 SO SEQUENCE 518 AA; 59507 MW; 2A52EC19028127F3 CRC64;

Query Match 100.0%; Score 151; DB 1; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 5; 7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDFDKYIVSASGDRITKIVW 29  
 DB 352 VGHRAAVNVDFDKYIVSASGDRITKIVW 380

## RESULT 2

FW1A\_HUMAN

AC Q9YKBL; Q9Y4C6; Q9P2S8; Q9P2S9; PRT; 542 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TrCP2).  
 GN FBXW1B OR FBW1B OR BTCP2 OR KIAA0696.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Gencliaralli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RT Pagano M.;  
 RL "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RC TISSUE=Petal lunc9;  
 RX MEDLINE=20160458; PubMed=10694485;  
 RA Kolje J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,  
 RT Katoh M.;  
 RL "Molecular cloning and genomic structure of the betarCP2 gene on  
 RL chromosome 5q35.1.";  
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RT Kotani H., Nomura N., Ohara O.;  
 RL "Prediction of the coding sequences of unidentified human genes. X.  
 RL The complete sequences of 100 new cDNA clones from brain which can  
 RL code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B AND C (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TR-ASP DOMAINS).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

DR EMBL; AF176022; AAF04528.1; -  
 DR EMBL; AB033279; BAA92329.1; -  
 DR EMBL; AB033280; BAA92330.1; -  
 DR EMBL; AB033281; BAA92331.1; -  
 DR EMBL; AB014596; BAA31671.1; ALT\_INT.  
 DR Genew; HGNC:13607; FBXW1B.  
 DR MIM; 605651; -

DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR Pfam; PF00646; F-box; 1.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR Prodom; PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS0181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW UBL conjugation pathway; Repeat: WD repeat; Alternative splicing.

FT DOMAIN 129 157 F-BOX.  
 FT REPEAT 238 275 WD 1.  
 FT REPEAT 278 315 WD 2.  
 FT REPEAT 318 355 WD 3.  
 FT REPEAT 361 398 WD 4.  
 FT REPEAT 401 440 WD 5.  
 FT REPEAT 442 478 WD 6.  
 FT REPEAT 490 527 WD 7.  
 FT VARSPIC 16 49 WD 7.  
 FT VARSPIC 16 48 MISSING (IN ISOFORM A).  
 FT VARSPIC 16 48 CSVPSRLMIGCANLVESGALSCLOSMPVRCVCL -> NTSV  
 MEDONEDSPKRTLM (IN ISOFORM B).  
 SO SEQUENCE 542 AA; 62090 MW; 7CD40087EFAA5C8A CRC64;

Query Match 100.0%; Score 151; DB 1; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 6e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDFDKYIVSASGDRITKIVW 29  
 DB 360 VGHRAAVNVDFDKYIVSASGDRITKIVW 388

## RESULT 3

FW1A\_HUMAN

AC Q9Y297; Q9Y213; STANDARD; PRT; 605 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TrCP)  
 DE (E3RSIKapab) (pikappaalpa-E3 receptor subunit).  
 GN FBXW1A OR FBW1A OR BTCP OR BTCP OR BTCP.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=99075339; PubMed=9859996;

RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
 RT "Identification of the receptor component of the IkappaBalpha-  
 RT ubiquitin ligase.";  
 RL Nature 396:590-594(1998).  
 RN (2)  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE-Lymphoid;  
 RX MEDLINE=98325370; PubMed=9660940;  
 RA Margottin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,  
 RA Thomas D., Strebel K., Benarous R.;  
 RT "A novel human WD protein, h-beta Trcp, that interacts with HIV-1 Vpu  
 RT connects CD4 to the ER degradation pathway through an F-box motif.";  
 RL Mol. Cell 1:565-574(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Cenciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99145464; PubMed=9990852;  
 RA Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,  
 RA Harper J.W.;  
 RT "The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically  
 RT with phosphorylated destruction motifs in I-kappa-B-alpha and  
 RT beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro.";  
 RL Genes Dev. 13:270-283(1999).  
 CC -1- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKR ALPHA  
 CC (PIKAPPAALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR  
 CC UBIQUITINATION AND DEGRADATION.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF101784; AAD08702.1; -  
 CC EMBL: Y14153; CAA74572.1; -  
 CC EMBL: AF129530; AAF04464.1; -  
 CC Genew: HNCN1144; BTRC.  
 DR MIM: 603482; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPOTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 KW DOMAIN 190 228  
 FT REPEAT 301 338 WD 1.  
 FT REPEAT 341 378 WD 2.  
 FT REPEAT 381 418 WD 3.  
 FT REPEAT 424 461 WD 4.  
 FT REPEAT 464 503 WD 5.  
 FT REPEAT 505 541 WD 6.

FT REPEAT 553 590 WD 7.  
 FT VARSPIC 17 52 MISSING (IN ISOFORM 2).  
 SO SEQUENCE 605 AA: 68866 MW: 4667F3B7EA00FD37 CRC64:  
 Query Match 100.0%; Score 151; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VGHRAAVNVVDFDDKIVYSASGDTIKVW 29  
 DB 423 VGHRAAVNVVDFDDKIVYSASGDTIKVW 451  
 RESULT 4  
 ID LI23\_CAEEL STANDARD; PRT; 665 AA.  
 AC Q09990; Q9GNN6;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein lin-23.  
 GN LIN-23 OR K10B2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, DEVELOPMENTAL STAGE, AND MUTAGENESIS.  
 RX MEDLINE=20515608; PubMed=11060233;  
 RA Kipreos E.T., Gohel S.P., Hedgecock E.M.;  
 RT "The Caenorhabditis elegans F-box/WD-repeat protein lin-23 functions  
 RT to limit cell division during development.";  
 RL Development 127:5071-5082(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN= Bristol N2;  
 RA Miller N.;  
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: Functions cell autonomously to negatively regulate cell  
 CC cycle progression. Required to restrain cell proliferation in  
 CC response to developmental cues. Probably recognizes and binds to  
 CC some proteins and promotes their ubiquitination and degradation  
 CC (by similarity).  
 CC -1- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex  
 CC (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).  
 CC -1- DEVELOPMENTAL STAGE: Highest levels in embryos and adults, lowest  
 CC levels in larvae. Maternal expression results in high zygotic  
 CC levels.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO X.LAEVIS PEXW1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF275253; AAG28037.1; -  
 CC EMBL: U28730; AAA68258.2; -  
 CC WormPep: K10B2.1; CE28600.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPOTEINBRPT.

DR ProDom: PD000018; WD40: 4.  
 DR SMART: SM00256; FBOX: 1.  
 DR SMART: SM00320; WD40: 7.  
 DR PROSITE: PS0181; FBOX: 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1: 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2: 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION: 1.  
 KW Developmental Protein; Cell cycle; Cell division;  
 Ubl conjugation pathway; Repeat; WD repeat.  
 FT DOMAIN 81 127 F-BOX.  
 FT REPEAT 220 257 WD 1.  
 FT REPEAT 260 299 WD 2.  
 FT REPEAT 301 337 WD 3.  
 FT REPEAT 343 380 WD 4.  
 FT REPEAT 383 420 WD 5.  
 FT REPEAT 423 460 WD 6.  
 FT REPEAT 472 509 WD 7.  
 FT DOMAIN 525 582 ALA-RICH.  
 FT MUTAGEN 441 441 G->R: IN LIN-32(RH293).  
 SQ SEQUENCE 665 AA; 75916 MW; BF3E9AF51F12ECC CRC64;

Query Match 98.0%; Score 148; DB 1; Length 665;  
 Best Local Similarity 96.6%; Pred. No. 2.1e-14;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDPDDKYIVSASGDRITKRW 29  
 |||||  
 DB 342 VGHRAAVNVDPDDRYIVSASGDRITKRW 370

## RESULT 5

SC02\_NEUCR

ID SC02\_NEUCR STANDARD; PRT; 650 AA.

AC Q01277;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sulfur controller-2 (SCON2).  
 GN SCON-2.  
 OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=74-OR23-1A;  
 RX MEDLINE=95241499; PubMed=7724564;  
 RA Kumar A., Paletta J.V.;  
 RT "The sulfur controller-2 negative regulatory gene of Neurospora  
 crassa encodes a protein with beta-transducin repeats.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).

CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.  
 CC -1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT  
 PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: U1251; AA68968.1;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001880; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 4.

DR SMART: SM00256; FBOX: 1.  
 DR SMART: SM00320; WD40: 7.  
 DR PROSITE: PS0181; FBOX: 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1: 2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2: 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION: 1.  
 KW Transcription regulation; Repeat; WD repeat.  
 FT DOMAIN 124 170 F-BOX.  
 FT REPEAT 292 320 WD 1.  
 FT REPEAT 332 360 WD 2.  
 FT REPEAT 372 400 WD 3.  
 FT REPEAT 411 441 WD 4.  
 FT REPEAT 453 488 WD 5.  
 FT REPEAT 528 564 WD 6.  
 FT REPEAT 576 604 WD 7.  
 FT REPEAT 616 644 WD 8.  
 SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 56.3%; Score 85; DB 1; Length 650;  
 Best Local Similarity 51.7%; Pred. No. 5.7e-05;  
 Matches 15; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDPDDKYIVSASGDRITKRW 29  
 |||||  
 DB 331 VGHAGIRALQFDKSLISGLDHTTKRW 359

## RESULT 6

SC0B\_EMENT

ID SC0B\_EMENT STANDARD; PRT; 678 AA.

AC Q00659;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sulfur metabolite repression control protein.  
 GN SCONB OR MAP1.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Natorff R.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE  
 REPRESSION.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT  
 PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: U21220; AAC15905.1;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX: 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS0181; FBOX: 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1: 4.  
 DR PROSITE: PS50082; WD\_REPEATS\_2: 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION: 1.  
 KW Transcription regulation; Repeat; WD repeat.



```

FT DOMAIN 178 224 F-BOX.
FT REPEAT 347 375 WD 1.
FT REPEAT 387 415 WD 2.
FT REPEAT 427 455 WD 3.
FT REPEAT 466 496 WD 4.
FT REPEAT 508 543 WD 5.
FT REPEAT 553 595 WD 6.
FT REPEAT 607 635 WD 7.
FT REPEAT 647 675 WD 8.
SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;

Query Match 55.0%; Score 83; DB 1; Length 678;
Best Local Similarity 50.0%; Pred. No. 0.00012;
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 GHRAAVNVDFDDKYIVASGDRITIKV 29
DB 387 GHESGRCLQFPDDTKLIGSMDRTIKV 414

RESULT 7
KMAH.DICDI STANDARD; PRT; 1146 AA.
AC P42527;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Myosin heavy chain kinase A (EC 2.7.1.129) (MHCK A).
GN MHKA OR MHCKA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=4689;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=AX3;
RX MEDLINE=95122486; PubMed=7822274;
RT Futey L.M., Medley O.G., Cote G.P., Egelhoff T.T.;
RT "Structural analysis of myosin heavy chain kinase A from
RT Dictyostelium. Evidence for a highly divergent protein kinase domain,
RT an amino-terminal coiled-coil domain, and a domain homologous to the
RT beta-subunit of heterotrimeric G proteins."
RT J. Biol. Chem. 270:523-529(1995).
RN [2]
RP CHARACTERIZATION OF THE CATALYTIC DOMAIN.
RC STRAIN=AX3;
RX MEDLINE=97207233; PubMed=9054368;
RA Cote G.P., Luo X., Murphy M.B., Egelhoff T.T.;
RT "Mapping of the novel protein kinase catalytic domain of
RT Dictyostelium myosin II heavy chain kinase A."
RT J. Biol. Chem. 272:6846-6849(1997).
CC -1- FUNCTION: PHOSPHORYLATES THREONINE IN THE C-TERMINAL TAIL REGION
CC OF MYOSIN II HEAVY CHAIN. THIS PHOSPHORYLATION IS CRITICAL IN
CC REGULATING THE ASSEMBLY AND DISASSEMBLY OF MYOSIN II FILAMENT.
CC REGQUIRES AUTOPHOSPHORYLATION FOR ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP + [myosin heavy-chain] = ADP + [myosin
CC heavy-chain] phosphate.
CC -1- COFACTOR: MAGNESIUM OR MANGANESE.
CC -1- SUBUNIT: OLIGOMER.
CC -1- DOMAIN: CONSISTS OF AN N-TERMINAL DOMAIN WITH PROBABLE COILED COIL
CC STRUCTURE, A CENTRAL NONREPEPTITIVE CATALYTIC DOMAIN, AND A C-
CC TERMINAL DOMAIN WITH SEVEN WD REPEATS.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE MHCK / EF-2 PROTEIN KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: U16856; AAA66070.1;
DR DictyDb; ID01086; mhka.
DR InterPro; IPR004166; MHCK_EF2_kinase.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF02816; MHCK_EF2_kinase; 1.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00676; WD_REPEATS_1; 4.
DR PROSITE; PS00082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW WD repeat; Phosphorylation; Coiled coil.
FT DOMAIN 100 120 COILED COIL (POTENTIAL).
FT DOMAIN 144 148 POLY-GLN.
FT DOMAIN 175 181 POLY-GLY.
FT DOMAIN 187 241 COILED COIL (POTENTIAL).
FT DOMAIN 297 502 COILED COIL (POTENTIAL).
FT DOMAIN 345 348 POLY-SER.
FT DOMAIN 438 441 POLY-LEU.
FT DOMAIN 500 551 PSEUDOSUBSTRATE/AUTOINHIBITORY DOMAIN
(POTENTIAL).
FT DOMAIN 552 852 CATALYTIC.
FT NP_BIND 778 783 ATP (POTENTIAL).
FT REPEAT 867 897 WD 1.
FT REPEAT 910 938 WD 2.
FT REPEAT 952 980 WD 3.
FT REPEAT 993 1021 WD 4.
FT REPEAT 1033 1061 WD 5.
FT REPEAT 1073 1101 WD 6.
FT REPEAT 1114 1142 WD 7.
SQ SEQUENCE 1146 AA; 128945 MW; 98D831F7948B5573 CRC64;

Query Match 54.3%; Score 82; DB 1; Length 1146;
Best Local Similarity 55.2%; Pred. No. 0.00031;
Matches 16; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 VGHRAAVNVDFDDKYIVASGDRITIKV 29
DB 951 VGHTRGVNCVAVNEKYLFCSDYDKTIKV 979

RESULT 8
PWP2.CAEEL STANDARD; PRT; 910 AA.
ID PWP2.CAEEL
AC P91341;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Periodic tryptophan protein 2 homolog.
DE F5F8.3.
OS Caenorhabditis elegans.
GN Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Gattung S., Scheet P., Kemp K.;
RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 14 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE PWP2 FAMILY OF WD-REPEAT PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U08047; AAB37807.2; -  
 WormRep; F55F8.3; CE30342.  
 InterPro: IPR001680; WD40.  
 Pfam: PF00400; WD40; 9.  
 PRINTS: PR00320; GPROTEINRPT.  
 ProDom: PD000018; WD40; 1.  
 SMART: SM00320; WD40; 12.  
 PROSITE: PS00678; WD\_REPEATS\_1; 1.  
 PROSITE: PS50082; WD\_REPEATS\_2; 3.  
 PROSITE: PS50294; WD\_REPEATS\_REGION; 2.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 12 50 WD 1.  
 FT REPEAT 93 92 WD 2.  
 FT REPEAT 94 134 WD 3.  
 FT REPEAT 144 183 WD 4.  
 FT REPEAT 188 227 WD 5.  
 FT REPEAT 271 310 WD 6.  
 FT REPEAT 313 355 WD 7.  
 FT REPEAT 358 397 WD 8.  
 FT REPEAT 400 439 WD 9.  
 FT REPEAT 443 485 WD 10.  
 FT REPEAT 486 523 WD 11.  
 FT REPEAT 525 563 WD 12.  
 FT REPEAT 586 625 WD 13.  
 FT REPEAT 688 728 WD 14.  
 SQ SEQUENCE 910 AA; 102197 MW; CC3C7DA636F131E1 CRC64;

Query Match 53.6%; Score 81; DB 1; Length 910;  
 Best Local Similarity 50.0%; Pred. NO. 0.00033;  
 Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRAAVNVDFDDKYIVSASGDRITKRW 29  
 II::: :I :IIIIIIIIII  
 DB 486 GHESAISSIDIGHNIHVGSGMDRTIKW 513

RESULT 9  
 KMBL\_DICDI STANDARD: PRT; 732 AA.  
 AC P90648;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Mysin heavy chain kinase B (EC 2.7.1.129) (MHCK B).  
 GN MHCKB OR MHCKB.  
 OS Dictyostelium discoideum (Slime mold).  
 CC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX3;  
 RA MEDLINE=97277316; PubMed=9115238;  
 RA Clancy C.E., Mendoza M.G., Naismith T.V., Kolman M.F., Egelhoff T.T.;  
 RT "Identification of a protein kinase from Dictyostelium with homology  
 RT to the novel catalytic domain of myosin heavy chain kinase A.";  
 RL J. Biol. Chem. 272:11812-11815(1997).  
 CC -1- FUNCTION: PHOSPHORYLATES THREONINE IN THE C-TERMINAL TAIL REGION  
 CC OF MYOSIN II HEAVY CHAIN. THIS PHOSPHORYLATION IS CRITICAL IN  
 CC REGULATING THE ASSEMBLY AND DISASSEMBLY OF MYOSIN II FILAMENT.  
 CC -1- CATALYTIC ACTIVITY: ATP + [myosin heavy-chain] = ADP + [myosin  
 CC heavy-chain] phosphate.  
 CC -1- DOMAIN: CONSISTS OF AN N-TERMINAL DOMAIN, A CENTRAL NONREPEATITIVE  
 CC CATALYTIC DOMAIN, AND A C-TERMINAL DOMAIN WITH SEVEN WD REPEATS.  
 CC -1- SIMILARITY: BELONGS TO THE MHCK / EF-2 PROTEIN KINASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial  
 entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U90946; AAB50136.1; -  
 DictyDB; DD01087; mhck.  
 InterPro: IPR004166; MHCK\_EF2\_kinase.  
 InterPro: IPR001680; WD40.  
 Pfam: PF00400; WD40; 7.  
 PRINTS: PF02816; MHCK\_EF2\_kinase; 1.  
 ProDom: PD000018; WD40; 6.  
 SMART: SM00320; WD40; 7.  
 PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
 WD repeat.  
 FT NP-BIND 298 303 ATP (POTENTIAL).  
 FT DOMAIN 355 381 POLY-ASN.  
 FT REPEAT 458 486 WD 1.  
 FT REPEAT 500 528 WD 2.  
 FT REPEAT 540 568 WD 3.  
 FT REPEAT 580 608 WD 4.  
 FT REPEAT 620 648 WD 5.  
 FT REPEAT 660 688 WD 6.  
 FT REPEAT 700 730 WD 7.  
 SQ SEQUENCE 732 AA; 83167 MW; A7233C4BD56D4088 CRC64;

Query Match 53.0%; Score 80; DB 1; Length 732;  
 Best Local Similarity 53.6%; Pred. NO. 0.00037;  
 Matches 15; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 2 GHRAAVNVDFDDKYIVSASGDRITKRW 29  
 II::: :I :IIIIIIIIII  
 DB 540 GHDKPVHTVLNDKXILFSGSGDRTIKW 567

RESULT 10  
 YEP6\_YEAST STANDARD: PRT; 185 AA.  
 AC P40042;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 20.7 kDa Trp-Asp repeats containing protein in IC11-NOT4  
 DE intergenic region.  
 GN YER066W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
 RA Chung E., Duncan M., Guzman E., Hartzel G., Hunkeler-Smith S.,  
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Natsath A., Norgren R., Oelner P., Oh C.,  
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
 RA Taylor P., Wei Y., Yellon M., Botstein D., Davis R.W.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 3 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U18813; AAB64602.1; -  
 DR SGD: S0000868; YER066W.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 3.  
 DR ProDom: PD000018; WD40; 1.  
 DR SMART: SM00320; WD40; 2.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 9 48 WD 1.  
 FT REPEAT 71 108 WD 2.  
 FT REPEAT 111 148 WD 3.  
 SQ SEQUENCE 185 AA; 20686 MW; A07DCEFB3CA30399 CRC64;  
 Query Match 51.0%; Score 77; DB 1; Length 185;  
 Best Local Similarity 50.0%; Pred. No. 0.00022;  
 Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 QY 2 GHRRAVNVDFDVKYIVSASGDKTIKW 29  
 DB 71 GHRALVGLGLSDKFLVSASVDSIRCM 98  
 RESULT 11  
 GBLP\_YEAST STANDARD; PRT; 319 AA.  
 ID GBLP\_YEAST  
 AC P38011;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 16-OCT-1996 (Rel. 34, Last sequence update)  
 DE Guanine nucleotide-binding protein beta subunit-1-like protein.  
 GN ASCL OR YMR16C OR YMR718.15C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C / AB972;  
 RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 54-62.  
 RC STRAIN=5288C.  
 RA Garrels J.I., Fletcher B., Kobayashi R., Latter G.I., Schwender B.,  
 RT Voipe T., Warner J.R., McLaughlin C.S.;  
 RT "Protein identifications for a Saccharomyces cerevisiae protein  
 database";  
 RT Electrophoresis 15:1466-1486(1994).  
 RN [3]  
 RP SEQUENCE OF 63-70 AND 138-149.  
 RC STRAIN=ATCC 44827 / SK02N;  
 RA MEDLINE=97190279; PubMed=9038161;  
 RA Norbeck J., Blomberg A.;  
 RT "Metabolic and regulatory changes associated with growth of  
 Saccharomyces cerevisiae in 1.4 M NaCl. Evidence for osmotic  
 induction of glycerol dissimilation via the dihydroxyacetone  
 pathway";  
 RT J. Biol. Chem. 272:5544-5554(1997).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: 249702; CAAB9754.1; -  
 DR SGD: S0004722; ASCL.

DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 3.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT REPEAT 15 46 WD 1.  
 FT REPEAT 63 93 WD 2.  
 FT REPEAT 105 135 WD 3.  
 FT REPEAT 146 182 WD 4.  
 FT REPEAT 194 224 WD 5.  
 FT REPEAT 235 264 WD 6.  
 FT REPEAT 284 314 WD 7.  
 SQ SEQUENCE 319 AA; 34805 MW; B9A5DA959A3C54FB CRC64;  
 Query Match 51.0%; Score 77; DB 1; Length 319;  
 Best Local Similarity 54.8%; Pred. No. 0.0004;  
 Matches 17; Conservative 4; Mismatches 8; Indels 2; Gaps 1;  
 QY 1 GHRRAVNVDFDVKYIVSASGDKTIKW 29  
 DB 104 VGHKSDVMSVDIDKKASMTISGRDRTIKW 134  
 RESULT 12  
 YCWM\_YEAST STANDARD; PRT; 515 AA.  
 ID YCWM\_YEAST  
 AC P25382;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 57.0 kDa Trp-Asp repeats containing protein in CPR4-SSK22  
 DE Intergenic region.  
 GN YCR072C OR YCR72C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ballesta J.P.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,  
 RA Sanz E.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP REVISIONS.  
 RA Jimenez A.;  
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: X59720; CAA42270.1; -  
 DR PIR: S19487; S19487.  
 DR SGD: S0000668; YCR072C.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 8.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 6.  
 DR SMART: SM00320; WD40; 8.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 141 172 WD 1.

FT REPEAT 184 214 WD 2.  
 FT REPEAT 227 264 WD 3.  
 FT REPEAT 276 305 WD 4.  
 FT REPEAT 318 387 WD 5.  
 FT REPEAT 400 430 WD 6.  
 FT REPEAT 442 472 WD 7.  
 FT REPEAT 484 514 WD 8.  
 SQ SEQUENCE 515 AA; 57040 MW; DFE03B79BF8530A CRC64;

Query Match  
 Best Local Similarity 50.3%; Score 76; DB 1; Length 515;  
 Matches 17; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 2 GHRAAVNVDF--DDKTVASGDRITIKW 29  
 DB 400 GHOKLVNHVAFSPDGRIYASAFDNGSIKW 429

RESULT 13  
 YRIO\_CAEEL STANDARD; PRT; 571 AA.  
 AC P90794;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DE Hypothetical 65.2 kDa Trp-Asp repeats containing protein D2030.9 in chromosome 1.  
 GN D2030.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 ON NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE LEC1AB FAMILY OF WD-REPEAT PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z73906; CAA98121.1;  
 DR WormPep; D2030.9; CE09086.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 5.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR ProDom; PD000018; WD40; 1.  
 DR SMART; SM00320; WD40; 5.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; FALSE\_NEG.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 3.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 309 340 WD 1.  
 FT REPEAT 357 387 WD 2.  
 FT REPEAT 439 470 WD 3.  
 FT REPEAT 482 512 WD 4.  
 SQ SEQUENCE 571 AA; 65170 MW; 521932753C6F8FC4 CRC64;

Query Match  
 Best Local Similarity 49.7%; Score 75; DB 1; Length 571;  
 Matches 16; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

QY 2 GHRAAVNVDF--DDKTVASGDRITIKW 29  
 DB 357 GHRCGVTHVDSRQDERYLNSKDDITIKW 386

RESULT 14  
 SE10\_CAEEL STANDARD; PRT; 579 AA.  
 AC Q93794;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE sel-10 protein.  
 GN sel-10 OR F55H12.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 ON NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP REVISIONS.  
 RA Jones S.J.M.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z79757; CAB02129.1;  
 DR WormPep; F55H12.3; CE16120.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR Pfam; PF00646; F-box; 1.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR ProDom; PD000018; WD40; 5.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT DOMAIN 113 159 F-BOX.  
 FT REPEAT 245 274 WD 1.  
 FT REPEAT 286 316 WD 2.  
 FT REPEAT 328 356 WD 3.  
 FT REPEAT 368 396 WD 4.  
 FT REPEAT 408 438 WD 5.  
 FT REPEAT 453 481 WD 6.  
 FT REPEAT 493 522 WD 7.  
 SQ SEQUENCE 579 AA; 64275 MW; 0647245AF5964663 CRC64;

Query Match  
 Best Local Similarity 49.7%; Score 75; DB 1; Length 579;  
 Matches 14; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 GHRAAVNVDFDDKTVASGDRITIKW 29  
 DB 368 GHHAAYRCVQFDCTTVSGGDTYKWI 395

RESULT 15  
 COPP\_DROME STANDARD; PRT; 914 AA.  
 ID COPP\_DROME  
 AC 062621;  
 DT 16-OCT-2001 (Rel. 40, Created)

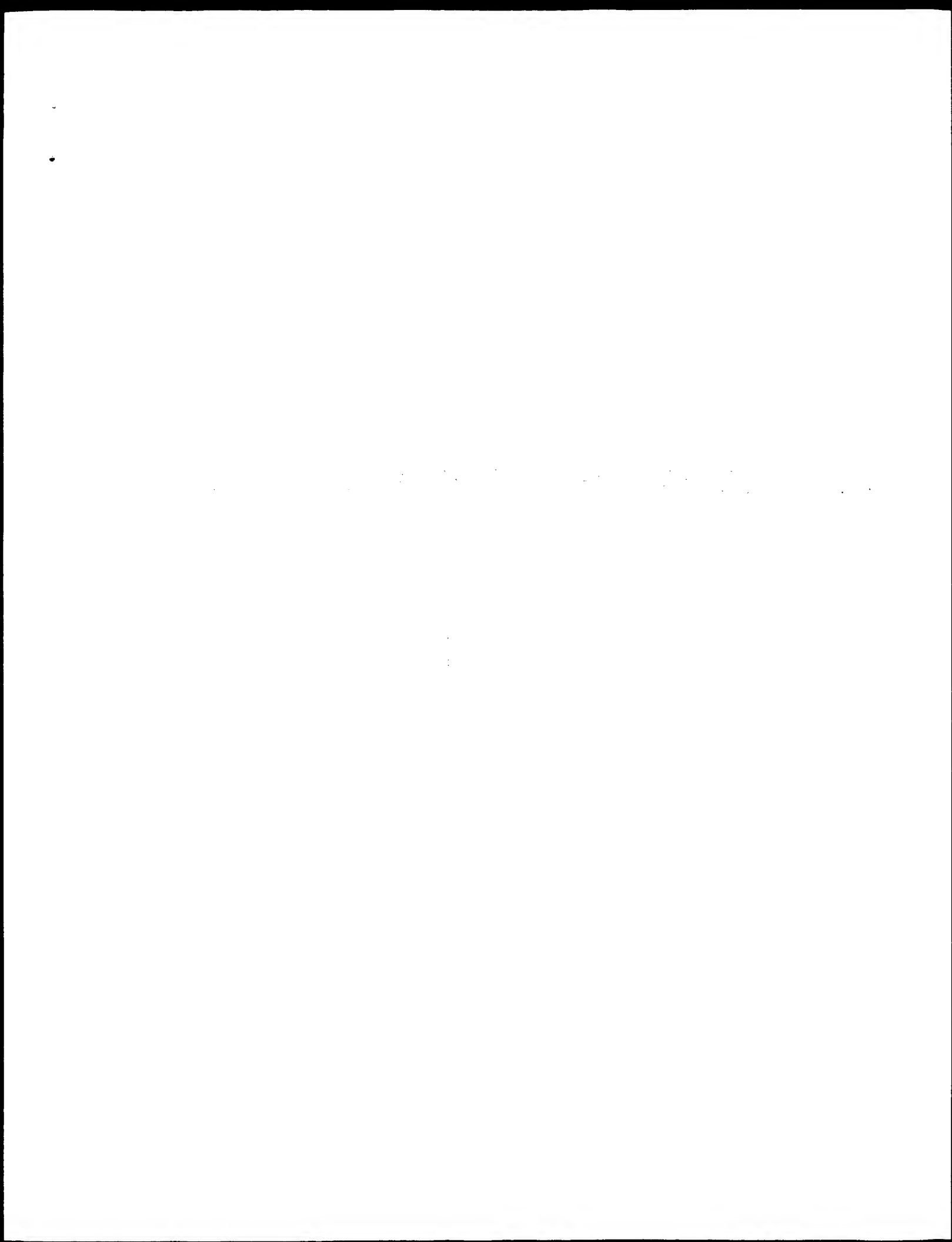
```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coatomer beta' subunit (Beta'-coat protein) (Beta'-COP).
GN BETA'-COP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Merdes G., Heid H.W., Mechler B.M.;
RT "Cloning and characterization of the Drosophila coatomer subunit
RT beta'";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE COATOMER IS A CYTOSOLIC PROTEIN COMPLEX THAT BINDS
CC TO DILYSINE MOTIFS AND REVERSIBLY ASSOCIATES WITH GOLGI NON-
CC CLATHRIN-COATED VESICLES, WHICH FURTHER MEDIATE BIOSYNTHETIC
CC PROTEIN TRANSPORT FROM THE ER, VIA THE GOLGI UP TO THE TRANS GOLGI
CC NETWORK. COATOMER COMPLEX IS REQUIRED FOR BUDDING FROM GOLGI
CC MEMBRANES, AND IS ESSENTIAL FOR THE RETROGRADE GOLGI-TO-ER
CC TRANSPORT OF DILYSINE-TAGGED PROTEINS (BY SIMILARITY).
CC -!- SUBUNIT: OLIGOMERIC COMPLEX THAT CONSIST OF AT LEAST THE ALPHA,
CC BETA, BETA', GAMMA, DELTA, EPSILON AND ZETA SUBUNITS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: THE COATOMER IS CYTOPLASMIC OR POLYMERIZED
CC ON THE CYTOPLASMIC SIDE OF THE GOLGI, AS WELL AS ON THE
CC VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AJ006524; CAA07085.1; -.
DR EMBL; AJ006523; CAA07084.1; -.
DR FLYbase; FBgn0025724; Beta'-COP.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 4.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
DR PROSITE; PS50294; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Transport; Protein transport; Golgi stack; Membrane; Repeat;
KW WD repeat; Endoplasmic reticulum.
FT REPEAT 13 43 WD 1.
FT REPEAT 55 85 WD 2.
FT REPEAT 97 127 WD 3.
FT REPEAT 140 171 WD 4.
FT REPEAT 183 215 WD 5.
FT REPEAT 227 257 WD 6.
SO SEQUENCE 914 AA; 102697 MW; C7AA25FAE5426695 CRC64;

Query Match 49.7%; Score 75; DB 1; Length 914;
Best Local Similarity 40.6%; Pred. No. 0.0026;
Matches 13; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

```

QY 2 GHRAAVNVDF---DDKIYVSGDRTIKW 29  
 DB 183 GHEKGVNCVYRHGGDKPYLISGADDLVKIW 214  
 Search completed: April 10, 2003, 13:19:36  
 Job time : 3.08404 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 ; Search time 6.85262 Seconds  
(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_387\_415

Perfect score: 151  
Sequence: 1 VGHRAAVNVDFDKYIVSASGDRITKW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SPREMBL\_21:\*
- 2: sp.archaea:\*
- 3: sp.bacteria:\*
- 4: sp.fungi:\*
- 5: sp.human:\*
- 6: sp.invertebrate:\*
- 7: sp.mammal:\*
- 8: sp.mhc:\*
- 9: sp.organelle:\*
- 10: sp.phage:\*
- 11: sp.plant:\*
- 12: sp\_rodent:\*
- 13: sp\_virus:\*
- 14: sp Vertebrate:\*
- 15: sp.unclassified:\*
- 16: sp.rvirus:\*
- 17: sp\_bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	100.0	122	13 080UN3	080un3 xenopus lae
2	151	100.0	252	11 0922C7	0922c7 mus musculi
3	151	100.0	569	11 09R1G7	09r1g7 mus musculi
4	151	100.0	569	11 0921S9	0921s9 mus musculi
5	151	100.0	569	11 0900U5	0900u5 mus musculi
6	147	97.4	510	5 044382	044382 drosophila
7	147	97.4	510	5 09VDE3	09vde3 drosophila
8	144	95.4	430	5 09H154	09h154 heterodera
9	133	88.1	563	11 0923H0	0923h0 mus musculi
10	83	55.0	423	16 08Y2Z3	08y2z3 anabaena sp
11	81	53.6	423	5 08SSJ1	08ssj1 encephalito
12	79	52.3	1376	3 08X1P2	08x1p2 podospira a
13	79	52.3	1551	16 08YMU3	08ymu3 anabaena sp
14	78	51.7	1326	5 09VZF4	09vzf4 drosophila
15	78	51.7	1747	16 08Z0Z0	08z0z0 anabaena sp
16	76	50.3	513	11 09D1I8	09d1i8 mus musculi

17	75	49.7	196	13 080GU0	080gu0 melanochrom
18	75	49.7	196	13 080GR9	080gr9 astatotiliap
19	75	49.7	196	13 080GR8	080gr8 labidochrom
20	75	49.7	196	13 080GR7	080gr7 trophus du
21	75	49.7	196	13 080GR6	080gr6 spathodus e
22	75	49.7	196	13 080GR5	080gr5 cyprichromi
23	75	49.7	196	13 080GR4	080gr4 xenotiliapia
24	75	49.7	196	13 080GR3	080gr3 neolampiroi
25	75	49.7	196	13 080GR2	080gr2 altolampiroi
26	75	49.7	196	13 080GR1	080gr1 steatocranu
27	75	49.7	196	13 080GR0	080gr0 tilapia but
28	75	49.7	196	13 080GS9	080gs9 oreochromis
29	75	49.7	196	13 080GS8	080gs8 mus musculi
30	75	49.7	503	11 09CY96	09cy96 homo sapien
31	75	49.7	513	4 09B059	09b059 homo sapien
32	75	49.7	513	4 09NVD1	09nvdi ratus norv
33	75	49.7	513	11 099M63	099m63 ratus norv
34	75	49.7	513	11 09UJ70	09uj70 mus musculi
35	75	49.7	513	5 09U3K9	09u3k9 caenorhabdi
36	75	49.7	585	5 095ZT0	095zt0 caenorhabdi
37	75	49.7	587	5 044083	044083 caenorhabdi
38	75	49.7	860	5 09VJ20	09vj20 drosophila
39	74	49.0	690	5 061585	061585 strongyloce
40	74	49.0	1189	16 08YTD1	08ytd1 anabaena sp
41	74	49.0	1711	16 08Z0I9	08z0i9 anabaena sp
42	73	48.3	326	10 09LV28	09lv28 arabidopsis
43	73	48.3	326	10 09C4Z6	09c4z6 arabidopsis
44	73	48.3	849	5 09U5J8	09u5j8 trypanosoma
45	72	47.7	689	10 093ZT0	093zt0 arabidopsis

ALIGNMENTS

RESULT 1  
ID 080UN3 PRELIMINARY; PRT; 122 AA.  
AC 080UN3;  
DT 01-MAR-2002 (TRENBLREL. 20, Created)  
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
DE Beta-Trcp protein (Fragment).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Carnevali F.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ballarino M.;  
RT "Analisi strutturale e funzionale del gene beta-Trcp in Xenopus laevis";  
RT Thesis (2001);  
RL Department of Genetics and Molecular Biology "Charles Darwin",  
RL University of Rome la Sapienza, Rome, Italy.  
DR EMBL: AJ428938; CAD21935.1; -  
DR InterPro: IPR01680; WD40.  
DR Pfam: PF00400; WD40.3  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR ProDom: PD000018; WD40.2.  
DR SMART: SM00320; WD40.3.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_2.  
DR PROSITE: PS0082; WD\_REPEATS\_2; 3.  
DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
FT NON\_TER 1  
FT 122  
SQ SEQUENCE 122 AA; 13682 MW; 365B6B39AC8F5387 CRC64;

```

Query Match          100.0%; Score 151; DB 13; Length 122;
Best Local Similarity 100.0%; Pred. No. 6,7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVVDFDDKTYIVSASGDRITKW 29
    |||||
DB 57 VGHRAAVNVVDFDDKTYIVSASGDRITKW 85

RESULT 2
0922C7 PRELIMINARY; PRT; 252 AA.
AC 0922C7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to f-box and WD-40 domain protein 1B (Fragment).
GN FBXW1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: BC008552; AA08552.1;
DR MGI: MGI:2144023; Fbxw1b.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 6.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 3.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT NON_TER
SQ SEQUENCE 252 AA; 28424 MW; F71737CBD7A9F75F CRC64;

Query Match          100.0%; Score 151; DB 11; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVVDFDDKTYIVSASGDRITKW 29
    |||||
DB 70 VGHRAAVNVVDFDDKTYIVSASGDRITKW 98

RESULT 3
09R1G7 PRELIMINARY; PRT; 569 AA.
AC 09R1G7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Beta-transducin repeat-containing protein.
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Winston J., Elledge S.J., Harper J.W.;
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF110396; AAD41025.1;
DR MGI: MGI:1338871; BTRC.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.

Query Match          100.0%; Score 151; DB 11; Length 569;
Best Local Similarity 100.0%; Pred. No. 4.5e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVVDFDDKTYIVSASGDRITKW 29
    |||||
DB 387 VGHRAAVNVVDFDDKTYIVSASGDRITKW 415

RESULT 5
09Q0U5 PRELIMINARY; PRT; 569 AA.
AC 09Q0U5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```

```

DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65047 MW; E6DDCAD28D51D9D CRC64;

Query Match          100.0%; Score 151; DB 11; Length 569;
Best Local Similarity 100.0%; Pred. No. 4.5e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVVDFDDKTYIVSASGDRITKW 29
    |||||
DB 387 VGHRAAVNVVDFDDKTYIVSASGDRITKW 415

RESULT 4
09Z1S9 PRELIMINARY; PRT; 569 AA.
AC 09Z1S9;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Beta-transducin repeat containing protein.
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99145465; PubMed=9990853;
RA Spencer E., Jiang J., Chen Z.-J.;
RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT S1mb/Beta-TrCP."
RL Genes Dev. 13:284-294(1999).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF112979; AAD04181.1;
DR MGI: MGI:1338871; BTRC.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65047 MW; BC7CA44815BED96 CRC64;

Query Match          100.0%; Score 151; DB 11; Length 569;
Best Local Similarity 100.0%; Pred. No. 4.5e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVVDFDDKTYIVSASGDRITKW 29
    |||||
DB 387 VGHRAAVNVVDFDDKTYIVSASGDRITKW 415

RESULT 5
09Q0U5 PRELIMINARY; PRT; 569 AA.
AC 09Q0U5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```



DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
 DE Ubiquitin ligase FWD1 (beta-transducin repeat containing protein)  
 GN (F-box-WD40 repeat protein 1).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9919275; PubMed=10097128;  
 RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,  
 RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,  
 RA Nakayama K., 1.  
 RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a  
 RT ubiquitin ligase Skp1/Cul1/F-box protein FWD1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99075339; PubMed=9859996;  
 RA Yaron A., Hatzubal A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,  
 RT "Identification of the receptor component of the IkappaBalpha-  
 RT ubiquitin ligase."  
 RL Nature 396:590-594(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.,  
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9129/SV;  
 RA Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,  
 RA Nakayama K., 1.  
 RT "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of  
 RT Caenorhabditis elegans SEL-10."  
 RL Genomics 78:214-223(2001).  
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF081887; A017755.1; -;  
 DR EMBL: AF099932; A008701.1; -;  
 DR EMBL: BC003989; A003989.1; -;  
 DR EMBL: AF391190; A040929.1; -;  
 DR EMBL: AF391178; A040929.1; JOINED.  
 DR EMBL: AF391179; A040929.1; JOINED.  
 DR EMBL: AF391180; A040929.1; JOINED.  
 DR EMBL: AF391181; A040929.1; JOINED.  
 DR EMBL: AF391182; A040929.1; JOINED.  
 DR EMBL: AF391183; A040929.1; JOINED.  
 DR EMBL: AF391184; A040929.1; JOINED.  
 DR EMBL: AF391185; A040929.1; JOINED.  
 DR EMBL: AF391186; A040929.1; JOINED.  
 DR EMBL: AF391187; A040929.1; JOINED.  
 DR EMBL: AF391188; A040929.1; JOINED.  
 DR EMBL: AF391189; A040929.1; JOINED.  
 DR MGI: MGI:1338871; Btcc.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS00682; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ligase; Repeat; WD repeat.  
 SQ SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Query Match 100.0%; Score 151; DB 11; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-14;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVDPDDKYIVSASGDRITKW 29  
 DB 387 VGHRAAVNVDPDDKYIVSASGDRITKW 415

RESULT 6  
 044382 PRELIMINARY; PRT; 510 AA.  
 AC 044382;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
 DE SLMB.  
 GN SLMB OR SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

Query Match 97.4%; Score 147; DB 5; Length 510;  
 Best Local Similarity 96.6%; Pred. No. 1.6e-13;

Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGHRAAVNVDPDDKYIVSASGDRITKW 29  
 DB 330 VGHRAAVNVDPDDKYIVSASGDRITKW 358

RESULT 7  
 09YDE3 PRELIMINARY; PRT; 510 AA.  
 AC 09YDE3;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
 DE SLMB protein (SLMB).  
 GN SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly)  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abdl J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Borchan M.R., Bouck J., Broksstein P., Brothier P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Daveport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Doonan K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hostin D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Turner C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissensbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of *Drosophila melanogaster*,"  
 RL Science 287:2185-2195(2000).  
 [12]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVARY, AND IMAGINAL DISC;  
 RX MEDLINE-20245299; PubMed-10781936;  
 RT Miletich I., Limbourg-Bouchon B.,  
 RT "Drosophila null limb clones transiently deregulate Hedgehog-  
 RT independent transcription of wingless in all limb discs, and induce  
 RT decapentaplegic transcription linked to imaginal disc regeneration,"  
 RL Mech. Dev. 93:15-26(2000).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF003733; AAF55853.1; -  
 DR EMBL: AF222924; AAF63214.1; -  
 DR EMBL: AF222923; AAF63213.1; -  
 DR Flybase: FBgn0023423; slmb.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SMO0256; FBOX; 1.  
 DR SMART: SMO0320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS50678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SO SEQUENCE 510 AA; 58552 MW; F4D5DF126F58A012 CRC64;

Query Match 97.4%; Score 147; DB 5; Length 510;  
 Best Local Similarity 96.6%; Pred. No. 1.6e-13;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 VGHRAAVNVDPDDKITYSASGDRITK 29  
 Db 330 VGHRAAVNVDPDDKITYSASGDRITK 358

RESULT 8  
 ID 09BJ54 PRELIMINARY; PRT: 430 AA.  
 AC 09BJ54;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Beta-transducin repeat-containing protein (fragment).  
 OS Heterodera glycines (Soybean cyst nematode).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;  
 OC Tylenchida; Heterodera; Heteroderidae; Heteroderinae; Heterodera.  
 NX NCBI\_TaxID=51029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kovaleva E.S., Yakovlev A.G., Masler E.P.,  
 RT "Plant parasitic nematode b-TTRCP,"  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF339101; AAK26376.1; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SMO0320; WD40; 7.  
 DR PROSITE: PS50678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER  
 SO SEQUENCE 430 AA; 47916 MW; 4ABC3F2DFE3A50B CRC64;  
 Query Match 95.4%; Score 144; DB 5; Length 430;  
 Best Local Similarity 96.6%; Pred. No. 3.6e-13;  
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 VGHRAAVNVDPDDKITYSASGDRITK 29  
 Db 211 VGHRAAVNVDPDDKITYSASGDRITK 239  
 RESULT 9  
 ID 0923H0 PRELIMINARY; PRT: 563 AA.  
 AC 0923H0;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE F-box/WD40 repeat-containing protein HOS.  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FVB/N;  
 RA Bhatia N., Herter J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;  
 RT "Mouse homolog of HOS (HMO5) is overexpressed in skin tumors and  
 RT implicated in constitutive activation of NF-kappaB,"  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AY038079; AAK72095.1; -  
 DR MGD: MGI:2144023; Fbxw1b.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRODOM: PD000018; WD40; 4.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS50678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.

DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 563 AA; 64741 MW; 9AB5623FF5E3496 CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 RAAVNVVDFDDKIVASGDRITIKW 29  
 DB 384 RAAVNVVDFDDKIVASGDRITIKW 409  
 RESULT 10  
 O8Y223 PRELIMINARY; PRT; 934 AA.  
 AC O8Y223;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE WD-40 repeat protein.  
 GN ALU0664.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2159285; PubMed=11759840; Kuritz T., Sasamoto S.,  
 RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Matanabe A., Iiguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL: AF003583; BAB72622.1; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 14.  
 DR PRINTS: PRO0320; GPROTEINBRPT.  
 DR SMART: SM00320; WD40; 14.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 13.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 934 AA; 104389 MW; D371F162040DFB3 CRC64;  
 Query Match  
 Best Local Similarity 55.0%; Score 83; DB 16; Length 934;  
 Matches 17; Conservative 4; Mismatches 7; Indels 2; Gaps 1;  
 OY 2 GHRAAVNVDF--DDKIVASGDRITIKW 29  
 DB 530 GHGQVNVSVTFSPDGKLIATASGDRITIKW 559  
 RESULT 11  
 O8SSJ1 PRELIMINARY; PRT; 423 AA.  
 AC O8SSJ1;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE WD-repeat protein similar to cell division control protein 4 / SEL10  
 DE protein.  
 GN ECU01\_1300.  
 OS Eucephalitozoon cuniculi.  
 OC Eukaryota; Microsporidia; Unikaryonidae; Eucephalitozoon.  
 OX NCBI\_TaxID=6035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 DE WD-repeat protein.  
 GN Genoscope;

RL submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Kallinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
 RA Prenstier G., Barbe V., Peyretillade E., Brotier P., Wincker P.,  
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivares C.P.;  
 RT "Genome sequence and gene compaction of the eukaryote parasite  
 Eucephalitozoon cuniculi.";  
 RL Nature 414:450-453(2001).  
 DR EMBL: AL391737; CAD25003.1; -  
 SQ SEQUENCE 423 AA; 47791 MW; 661672B1621620CF CRC64;  
 Query Match  
 Best Local Similarity 53.6%; Score 81; DB 5; Length 423;  
 Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
 OY 1 VGHRAAVNVVDFDDKIVASGDRITIKW 29  
 DB 151 VGHKGCVTFMFNDKIVSGSTDKTARIW 179  
 RESULT 12  
 O8X1P2 PRELIMINARY; PRT; 1376 AA.  
 AC O8X1P2;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Beta transducin-like protein HET-D2Y.  
 GN HET-D.  
 OS Podospira anserina.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OX Sordariales; Lasiosphaeriaceae; Podospira.  
 OX NCBI\_TaxID=5145;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Espagne E., Balhadere P., Bequeret J., Turcq B.;  
 RT "A new family of WD40 proteins implicated in vegetative  
 incompatibility; evidence for a major role of WD40 repeat domain in  
 the specificity of het-e and het-d genes.";  
 RL Submitted (Nov-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF323585; AAL37301.1; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 12.  
 DR PRINTS: PRO0320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 9.  
 DR SMART: SM00320; WD40; 12.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_11.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 11.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 SQ SEQUENCE 1376 AA; 152068 MW; 2191826A5C62D97 CRC64;  
 Query Match  
 Best Local Similarity 52.3%; Score 79; DB 3; Length 1376;  
 Matches 16; Conservative 4; Mismatches 8; Indels 2; Gaps 1;  
 OY 2 GHRAAVNVDF--DDKIVASGDRITIKW 29  
 DB 1123 GHGQVNVSVAFSPDSKVMASGSDRTIKIW 1152  
 RESULT 13  
 O8YMU3 PRELIMINARY; PRT; 1551 AA.  
 AC O8YMU3;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE WD-repeat protein.  
 GN ALU4834.

OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneke T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Matanabe A., Iritaguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genome sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL: AP003597; BAB76533.1;  
 DR InterPro: IPR000408; Reg\_chit\_condens.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 14.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR SMART: SM00320; WD40; 14.  
 DR PROSITE: PS00626; RC01\_2; UNKNOWN\_1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_9.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 14.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 1551 AA; 176680 MW; 8FEB62B3BA727317 CRC64;

Query Match 52.3%; Score 79; DB 16; Length 1551;  
 Best Local Similarity 46.7%; Pred. No. 0.01; Mismatches 14; Conservative 7; Indels 2; Gaps 1;

QY 2 GHRAAVNVDF-DDKYIVASGDRITKVV 29  
 Db 1484 GHRGAWSISIPDDYIATASDDRTIVRLW 1513

RESULT 14  
 Q9VZF4 PRELIMINARY; PRT; 1326 AA.  
 AC Q9VZF4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE CG15010 protein (L213222) (L2132221p).  
 GN AGO OR CG15010.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle G., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abrell J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borzova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskren D.R., Pacled J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svistskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye Y., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).

RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Abmayyan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E.,  
 RA Gonzalez M., Guarini H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunoo J., Pacled J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,  
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Abmayyan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarini H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacled J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AE003480; AAC22246.1;  
 DR EMBL: AY05401; AAL68231.1;  
 DR EMBL: AY05401; AAL68231.1;  
 DR FlyBase: FBgn0041171; ago.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 2.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 1326 AA; 141360 MW; 3F42C873CFA3027F CRC64;

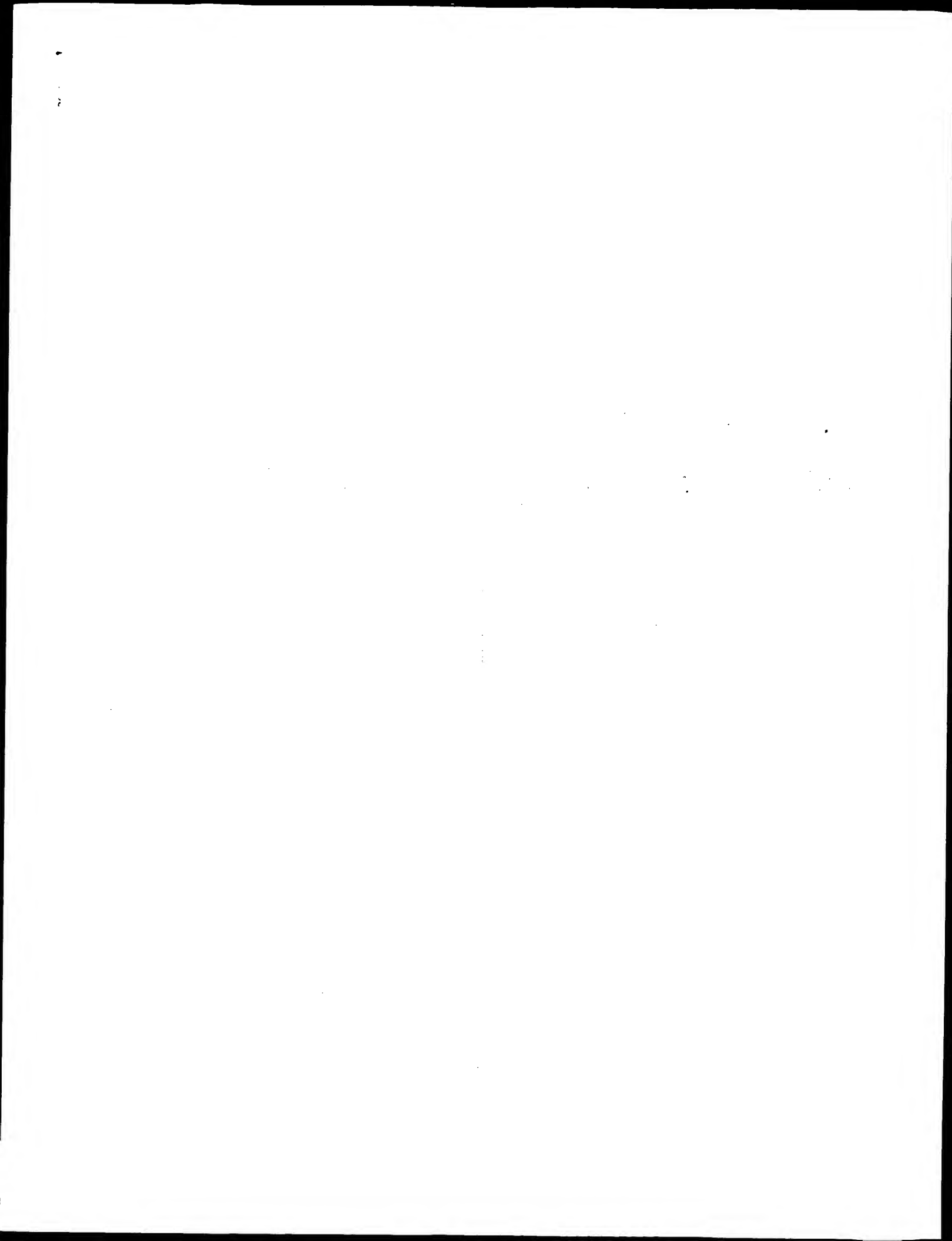
Query Match 51.7%; Score 78; DB 5; Length 1326;  
 Best Local Similarity 51.7%; Pred. No. 0.012;  
 Mismatches 15; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 VGRRAVNVDFDDKYIVASGDRITKVV 29  
 Db 1112 VGHAAVRCVQYDGLKLVGAYDYMKW 1140

RESULT 15  
 Q8Z020 PRELIMINARY; PRT; 1747 AA.  
 AC Q8Z020;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)  
DE WD-40 repeat protein.  
GN ALL0283.  
OS Anabaena sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Irliguchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
cyanobacterium Anabaena sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
DR EMBL: AP003581; BAB7807.1; -  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 14.  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR SMART: SM00320; WD40; 14.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 12.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Complete proteome.  
SQ SEQUENCE 1747 AA; 192610 MW; 27DC1CF2DCC8BDACA CRC64;  
  
OY 2 GHRAAVNVDF--DDKIYVSASGDRITIKW 29  
DB 1135 GHAAQVNAVSESPDGKVLASASDDRTYKWLW 1164  
  
Query Match 51.7%; Score 78; DB 16; Length 1747;  
Best Local Similarity 53.3%; Pred. No. 0.017;  
Matches 16; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

Search completed: April 11, 2003, 11:51:44  
Job time : 14.8526 secs



GenCore version 5.1.4-P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:20:44 ; Search time 8.93666 Seconds  
(without alignments)  
432.406 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_427\_455

Perfect score: 156  
Sequence: 1 NGRKRGICLCQYRDLRYVSSGSDNTIRLM 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	156	100.0	517	16 AAR85852	WD-40 domain-contg
2	156	100.0	542	21 AAY96696	Human E3 ubiquitin
3	156	100.0	542	21 AAM79127	Human protein SEQ
4	156	100.0	542	22 AAM40208	Human polypeptide
5	156	100.0	550	22 AAM41994	Human polypeptide
6	156	100.0	569	20 AAY24054	A human beta-trans
7	156	100.0	569	21 AAB12812	Mouse ubiquitin 11
8	156	100.0	569	21 AAB12813	Human beta-transdu
9	156	100.0	569	21 AAY96697	Human beta-TRCP.
10	156	100.0	569	21 AAY83041	F-box protein FBP-

11	156	100.0	569	21 AAY83250	F-box protein hbet
12	156	100.0	569	21 AAY83254	F-box protein FWD1
13	156	100.0	569	21 AAY44249	Human cell signa11
14	156	100.0	569	22 AAB48298	Human ZF11 protein
15	156	100.0	579	22 AAB78383	Human protein SEQ
16	156	100.0	590	22 AAM00847	Human bone marrow
17	156	100.0	605	22 AAM78582	Human bone marrow
18	156	100.0	608	22 AAM00960	Human protein SEQ
19	156	100.0	632	22 AAM78584	Human protein SEQ
20	156	100.0	654	22 AAM79566	Human protein SEQ
21	156	100.0	654	22 AAM79567	Human protein SEQ
22	156	100.0	654	22 AAM79568	Human protein SEQ
23	152	97.4	510	22 AAB59857	Drosophila melanog
24	150	96.2	29	16 AAR84939	Peptide r1 from a
25	87	55.8	34	16 AAR84935	Peptide r1 from a
26	83	53.2	1326	22 ABB67237	Drosophila melanog
27	83	53.2	1326	22 ABB67238	Drosophila melanog
28	83	53.2	1326	22 ABB70051	Drosophila melanog
29	82.5	52.9	28	16 AAR84936	Peptide r1 from a
30	80	51.3	694	21 AAY99661	Human GTPase assoc
31	79	50.6	357	23 ABB90392	Human polypeptide
32	79	50.6	438	21 AAG07498	Arabidopsis thailia
33	79	50.6	438	21 AAG53870	Human polypeptide
34	79	50.6	540	20 AAY22465	Human mammary sel-
35	79	50.6	540	20 AAY22468	Human mammary sel-
36	79	50.6	540	20 AAB59197	Human mammary sel-
37	79	50.6	540	22 AAB59200	Human mammary sel-
38	79	50.6	545	20 AAY22464	Human hippocampal
39	79	50.6	545	22 AAB59196	Human hippocampal
40	79	50.6	553	20 AAY22463	Human protein sequ
41	79	50.6	553	22 AAB93475	Human hippocampal
42	79	50.6	553	22 AAB59195	Human mammary sel-
43	79	50.6	559	22 AAY22467	Human mammary sel-
44	79	50.6	559	22 AAB59199	Human mammary sel-
45	79	50.6	589	20 AAY22466	Human mammary sel-

#### ALIGNMENTS

RESULT 1  
ID AAR85852 standard; peptide: 517 AA.  
XX AAR85852:  
AC AAR85852:  
XX 13-SEP-1996 (first entry)  
XX  
XX WD-40 domain-contg. beta-TRCP protein.  
DE  
XX  
XX WD40 repeat region; beta-transducin; protein-protein interaction; drug;  
KW intracellular signalling; protein kinase C; homology; motif; modulator;  
KW receptors of activated protein kinase; enzyme activity; isozyme; human.  
XX  
XX  
OS Synthetic.  
XX  
XX WO9521252-A2.  
XX  
XX PD 10-AUG-1995.  
XX  
XX PF 31-JAN-1995; 95MO-US01210.  
XX  
XX PR 01-FEB-1994; 9AUS-0190802.  
XX  
XX (STRD ) UNIV IELAND STANFORD JUNIOR.  
XX  
XX Mochly-Rosen D, Ron D;  
XX  
XX WPI: 1995-283772/37.  
XX  
XX New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the  
PT activity of a protein, eg. protein kinase C, which interacts with a  
PT protein contg. a WD-40 region.

XX Example 5; Page 80-82; 351pp; English.

PS  
XX  
CC Proteins AAR85851-92 are protein which contain at least one WD-40 (also  
CC called beta-transducing homologous) amino acid repeat motifs. The WD-40  
CC regions are involved in protein-protein interactions between proteins  
CC involved in intracellular signalling. An example of such an interaction  
CC is between protein kinase C and receptors of activated protein kinase  
CC (RACK), esp. RACK-1 (AAR85850). Proteins AAR85851-82 were isolated based  
CC on homology with beta-transducin, whereas proteins AAR85882-92 were  
CC isolated based on homology with the WD-40 consensus sequence (AAR85893).  
CC The proteins were used to construct the peptides AAR84928-R85063 and  
CC AAR85786-R85842. The peptides can be used to identify target proteins  
CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of  
CC proteins involved in protein-protein interaction and to screen for drugs  
CC that will affect protein-protein interaction involving WD-40 domains.

SQ Sequence 517 AA:

Query Match 100.0%; Score 156; DB 16; Length 517;  
Best Local Similarity 100.0%; Pred. No. 2, 2e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKGACIQLQYRDLVVGSSDNTIRLM 29  
Db 391 NGHKGACIQLQYRDLVVGSSDNTIRLM 419

RESULT 2  
ID AAY96696  
XX AAY96696 standard; Protein; 542 AA.  
XX  
AC AAY96696;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Human E3 ubiquitin ligase.  
XX  
KW E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
KW anti-inflammatory; immunosuppressive; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200034447-A2.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99WO-US29371.  
XX  
PR 10-DEC-1998; 98US-0210060.  
XX  
PA (SIGN-) SIGNAL PHARM INC.  
PA (YISS) YISSUM RES & DEV CO.  
XX  
PI Manning AM, Mercurio F, Amlt S, Ben-eriah Y, Davis M, Hatzubai A,  
PI Lavon I, Yaron A;  
XX  
DR WPI: 2000-431294/37.  
XX  
N-PSDB; AAK511228.  
XX  
PT polypeptide enhancing phosphorylated I-kappa-B ubiquitination useful for  
PT treating disorder associated with NF-kappa-B activation e.g. cancer,  
PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
PT variant  
XX  
PS Claim 1; Page 70-72; 77pp; English.  
XX  
XX This is human E3 ubiquitin ligase (E3), which is homologous to human  
CC beta-TrCP, an F-box/WD protein family member. E3 enhances ubiquitination  
CC of phosphorylated I-kappa-B, an inhibitor of nuclear factor  
CC kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the  
CC ubiquitin pathway is useful for identifying modulators of this process

CC for use in treating diseases associated with activation of NF-kappa-B. In  
CC vitro analysis suggests that deletion of the F-box results in a protein  
CC that functions as a dominant negative molecule in vivo. Transient  
CC over-expression of delta-beta-TrCP (a deletion mutant) inhibited the  
CC degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells,  
CC resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be  
CC used to screen for modulators of NF-kappa-B activity. E3 and beta-TrCP  
CC can be used to modulate NF-kappa-B to treat inflammatory diseases,  
CC autoimmune diseases, cancer and viral infections.

SQ Sequence 542 AA:

Query Match 100.0%; Score 156; DB 21; Length 542;  
Best Local Similarity 100.0%; Pred. No. 2, 3e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKGACIQLQYRDLVVGSSDNTIRLM 29  
Db 400 NGHKGACIQLQYRDLVVGSSDNTIRLM 428

RESULT 3  
ID AAM79127  
XX AAM79127 standard; Protein; 542 AA.  
XX  
AC AAM79127;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human protein SEQ ID NO 1789.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200157190-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US04098.  
XX  
PR 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-05060875.  
PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0654936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0693325.  
PR 30-NOV-2000; 2000US-0728422.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Dymnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang Y, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
DR WPI: 2001-476283/51.  
XX  
N-PSDB; AAK52260.  
XX  
PT Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
PS Claim 20; Page 4147-4148; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,



CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAK6020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

CC Sequence 542 AA:

Query Match 100.0%; Score 156; DB 22; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-16;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKRGIACTQYRDRLVSGSSDNTIRLM 29  
 |||||  
 Db 400 NGHKRGIACTQYRDRLVSGSSDNTIRLM 428

RESULT 4  
 AAM40208 standard; Protein; 542 AA.

XX AAM40208;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3353.

DE Human: neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX Homo sapiens.

XX W020015312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.  
 XX N-PSDB: AAI59364.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries -

XX Example 5: SEQ ID NO 3353; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 XX the encoded polypeptides (AAM38642-AAM42213) with neurotropic,  
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
 XX in gene therapy. A composition containing a polypeptide or polynucleotide  
 XX of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemia and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

CC Sequence 542 AA:

Query Match 100.0%; Score 156; DB 22; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-16;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKRGIACTQYRDRLVSGSSDNTIRLM 29  
 |||||  
 Db 400 NGHKRGIACTQYRDRLVSGSSDNTIRLM 428

RESULT 5  
 AAM41994 standard; Protein; 550 AA.

XX AAM41994;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6925.

DE Human: neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX Homo sapiens.

XX W020015312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.  
 XX N-PSDB: AAI61150.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries -

XX Example 2: SEQ ID NO 6925; 10078pp; English.  
 XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 XX the encoded polypeptides (AAM38642-AAM42213) with neurotropic,





Db 427 NGHKGIACTQYRDRLVVGSSDNTIRLM 455

RESULT 10

ID AAY83041 standard; Protein; 569 AA.

AC AAY83041;

DE 16-AUG-2000 (first entry)

DE F-box protein FBP-1.

KW F-box protein; FBP; diagnosis; treatment; screening; agonist;  
KW antagonist; proliferative disorder; differentiative disorder;  
KW breast cancer; prostate cancer; ovarian cancer; cancer;  
KW small cell lung carcinoma; immune disorder; cardiovascular disorder;  
KW inflammatory disorder; human.

OS Homo sapiens.

PN WO200012679-A1.

PD 09-MAR-2000.

PF 27-AUG-1999; 99WO-US19560.

PR 28-AUG-1998; 98US-0098355.

PR 03-FEB-1999; 99US-0118568.

PR 15-MAR-1999; 99US-0124449.

PA (UYNY ) UNIV NEW YORK STATE.

PI Chiaux DS, Pagano M, Latres E;

DR WPI; 2000-256635/22.

DR N-PSDB; AA293350.

PT Novel nucleic acid for screening compounds useful for treating

PT proliferative and differentiative disorders such as cancer and immune

PT disorders comprises sequences encoding ubiquitin ligases

PS Disclosure: Figure 3a; 245pp; English.

CC Nucleic acids encoding substrate-targeting subunits of ubiquitin  
CC ligases with F-box motifs (F-box proteins) are useful for diagnosis  
CC of proliferative and differentiated related disorders by measuring  
CC FBP gene expression. Cells expressing such proteins or  
CC their fragments are useful for screening compounds. The compounds  
CC are agonists or antagonists, which are useful for treating a  
CC proliferative or differentiative disorder in a mammal such as  
CC breast, ovarian and prostate cancer and small cell lung carcinoma  
CC and also major opportunistic infections, immune disorders,  
CC cardiovascular diseases and inflammatory disorders. FBP protein,  
CC analogs, derivatives and their subsequences, anti-FBP antibodies  
CC are also useful in diagnosis of the disorders.

Sequence 569 AA:

Query Match 100.0%; Score 156; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGHKGIACTQYRDRLVVGSSDNTIRLM 29

Db 427 NGHKGIACTQYRDRLVVGSSDNTIRLM 455

RESULT 11

ID AAY83250 standard; Protein; 569 AA.

AC AAY83250;

DE 16-AUG-2000 (first entry)

DE F-box protein hbetaTrcp.

KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;  
KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;  
KW oncoprotein; Huntington's disease; gene knockout; delivery systems;  
KW human.

OS Homo sapiens.

PN WO200022110-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US23705.

PR 09-OCT-1998; 98US-0103787.

PA (HARD ) HARVARD COLLEGE.

PI Zhou P, Howley P;

DR WPI; 2000-317970/27.

DR N-PSDB; AA293710.

PT Targeting degradation of polypeptide useful for treating cancer and  
PT other proliferative disorders, involves conjugating polypeptide with  
PT ubiquitin protein ligase or inhibiting ubiquitination using organic  
PT compound

PS Claim 9; Page 171; 185pp; English.

CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin  
CC ligases) which can be used for the targeted degradation of a target  
CC polypeptide in vivo. Targeted degradation is achieved by expressing  
CC the ubiquitin ligase in a cell linked to the interaction domain of  
CC the target polypeptide and thereby recruiting the target polypeptide  
CC to the ubiquitin ligase. Such methods are useful for decreasing or  
CC increasing the level of a target polypeptide and for creating and  
CC expressing a destabilized polypeptide which is subjected to SCF  
CC mediated proteolysis. Degrading any desired protein in a cell is  
CC useful for preventing or treating diseases caused by the presence of  
CC abnormal amount of the specific polypeptides, for drug discovery and  
CC for gene therapy. Diseases treated include cancer, by degradation of  
CC oncoproteins, Huntington's disease, other proliferative disorders and  
CC microbial infections. The method provides a quick and easy  
CC alternative to gene knockout technology. The target polypeptide can  
CC be degraded at all stages, or a specific stage, of development in the  
CC mature animal.

Sequence 569 AA:

Query Match 100.0%; Score 156; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGHKGIACTQYRDRLVVGSSDNTIRLM 29

Db 427 NGHKGIACTQYRDRLVVGSSDNTIRLM 455

RESULT 12

ID AAY83254 standard; Protein; 569 AA.

AC AAY83254;

DE 16-AUG-2000 (first entry)

DE F-box protein FWDp.

XX	ubiquitin ligase; SCF; F-box protein; targeted degradation;
KW	destabilisation; proteolysis; drug discovery; gene therapy; cancer;
KM	oncoprotein; Huntington's disease; gene knockout; delivery systems;
XX	mouse; ss.
OS	Mus musculus.
PN	MO200022110-A2.
PD	20-APR-2000.
XX	
PF	08-OCT-1999; 99WO-US23705.
XX	
PR	09-OCT-1998; 98US-0103787.
XX	
PA	(HARD ) HARVARD COLLEGE.
PI	Zhou P, Howley P;
XX	
DR	WPI; 2000-317970/27.
DR	N-PSDB; AA293714.
XX	
PT	Targeting degradation of polypeptide useful for treating cancer and
PT	other proliferative disorders, involves conjugating polypeptide with
Pt	ubiquitin protein ligase or inhibiting ubiquitination using organic
PT	compound
XX	
PS	Claim 9; Page 184-185; 185pp; English.
XX	
CC	The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
CC	ligases) which can be used for the targetted degradation of a target
CC	polypeptide in vivo. Targetted degradation is achieved by expressing
CC	the ubiquitin ligase in a cell linked to the interaction domain of
CC	the target polypeptide and thereby recruiting the target polypeptide
CC	to the ubiquitin ligase. Such methods are useful for decreasing or
CC	increasing the level of a target polypeptide and for creating and
CC	expressing a destabilized polypeptide which is subjected to SCF
CC	mediated proteolysis. Degrading any desired protein in a cell is
CC	useful for preventing or treating diseases caused by the presence of
CC	abnormal amount of the specific polypeptides, for drug discovery and
CC	for gene therapy. Diseases treated include cancer, by degradation of
CC	oncoproteins, Huntington's disease, other proliferative disorders and
CC	microbial infections. The method provides a quick and easy
CC	alternative to gene knockout technology. The target polypeptide can
CC	be degraded at all stages, or a specific stage, or development in the
CC	mature animal.
CC	
XX	
SO	Sequence 569 AA:
QY	
Query Match	100.0%; Score 156; DB 21; Length 569;
Best Local Similarity	100.0%; Pred. No. 2.4e-16;
Matches 29; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
OY	1 NGHKRCIACLOAYRDRLVSSSDNTIRLW 29 
Db	427 NGHKRGIAQLQYRDRLVSSSDNTIRLW 455
RESULT 13	
AA44249	
ID	AA44249 standard; Protein; 569 AA.
XX	
AC	AA44249;
XX	
DT	28-FEB-2000 (first entry)
XX	
DE	Human cell signalling protein-12.
XX	
KW	Cell signalling protein-12; CSIGP-12; cell proliferation;
KW	Inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;
KW	arteriosclerosis; Addison's disease; multiple sclerosis.
XX	

OS	Homo sapiens.	
XX		Location/Qualifiers
FH	Key	
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	19 /note= "Potential phosphorylation site"
FT	Modified-site	39 /note= "Potential phosphorylation site"
FT	Modified-site	91 /note= "Potential phosphorylation site"
FT	Modified-site	109 /note= "Potential phosphorylation site"
FT	Modified-site	162 /note= "Potential phosphorylation site"
FT	Modified-site	162 /note= "Potential phosphorylation site"
FT	Modified-site	266 /note= "Potential phosphorylation site"
FT	Modified-site	288 /note= "Potential phosphorylation site"
FT	Modified-site	328 /note= "Potential phosphorylation site"
FT	Modified-site	376 /note= "Potential phosphorylation site"
FT	Modified-site	381 /note= "Potential phosphorylation site"
FT	Modified-site	411 /note= "Potential phosphorylation site"
FT	Modified-site	418 /note= "Potential phosphorylation site"
FT	Modified-site	451 /note= "Potential phosphorylation site"
FT	Modified-site	514 /note= "Potential phosphorylation site"
FT	Modified-site	519 /note= "Potential phosphorylation site"
FT	Modified-site	535 /note= "Potential phosphorylation site"
FT	Modified-site	536 /note= "Potential phosphorylation site"
FT	Modified-site	17 /note= "Potential glycosylation site"
FT	Modified-site	77 /note= "Potential glycosylation site"
FT	Modified-site	416 /note= "Potential glycosylation site"
FT	Region	/label= Signature_sequence 320..334
FT	Region	/label= Signature_sequence 360..374
FT	Region	/label= Signature_sequence 403..417
FT	Region	/label= Signature_sequence 443..457
FT	Region	/label= Signature_sequence 483..497
FT	Region	/label= Signature_sequence 532..546
FT	Region	/label= Signature_sequence
PX	W09J58558-A2.	
PD	18-NOV-1999.	
PX	13-MAY-1999.	99MO-US10567.
PX	13-MAY-1998;	98US-0085343.
PR	26-AUG-1998;	98US-0098010.
XX	(INCY-) INCYTE PHARM INC.	
PX	Bandman O,	Lal P,
PI	Baugin MR,	Yang J;
XI		
DR	WP1; 2000-086432/07.	
DR	N-PDB; AAZ29233.	
XX		

PT Human cell signaling proteins useful for, e.g. diagnosing cell  
 PT proliferative and inflammatory disorders -  
 PS Claim 1; Page 77-78; 90pp; English.  
 CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded  
 CC by cDNA obtained from Incyte clone 3239149 of COLAUC01 library. It is  
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is  
 CC found to be homologous to beta-transducin repeats containing  
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as  
 CC hybridisation probe for detecting CSIGP related sequences or allelic  
 CC variants. Recombinant CSIGP can be produced in host cells by transforming  
 CC cells with genetically engineered vectors. Agonists or antagonists can be  
 CC used in the treatment of cell proliferative and inflammatory disorders  
 CC associated with decreased or increased CSIGP expression. CSIGP is used in  
 CC the diagnosis, prevention and treatment of cell proliferative disorders  
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.  
 SQ Sequence 569 AA:  
 Query Match 100.0%; Score 156; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 NGHKRGIACTQYRDRLVVGSSDNTIRLW 29  
 DB 427 NGHKRGIACTQYRDRLVVGSSDNTIRLW 455  
 RESULT 14  
 AAB48298  
 ID AAB48298 standard; protein; 569 AA.  
 AC AAB48298;  
 XX  
 DT 02-APR-2001 (first entry)  
 DE Human zfp1 protein.  
 XX  
 KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZP;  
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
 KW Bad; Bcl-2; tumour; cytostatic.  
 XX  
 OS Homo sapiens.  
 PA WO200075184-A1.  
 PN 14-DEC-2000.  
 PD  
 XX  
 PF 05-JUN-2000; 2000MO-US15449.  
 XX  
 PR 04-JUN-1999; 99US-0137494.  
 XX  
 PA (UYVA) UNIT VALE.  
 PI Zhang H, Tsvetkov LM, Kondo T;  
 XX WPI: 2001-061703/07.  
 DR N-PSDB; AAC84610.  
 XX  
 PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
 PT involves altering levels of proteins such as S-phase kinase associated  
 PT proteins 1, 2 and cullin/CDC53 proteins -  
 PS Claim 3; Page 130-132; 162pp; English.  
 CC The invention relates to methods of altering the polypeptide levels in a  
 CC cell, using proteins selected from S-phase kinase associated proteins 1  
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
 CC cullin/CDC53 family of proteins). The method is useful for altering the  
 CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for

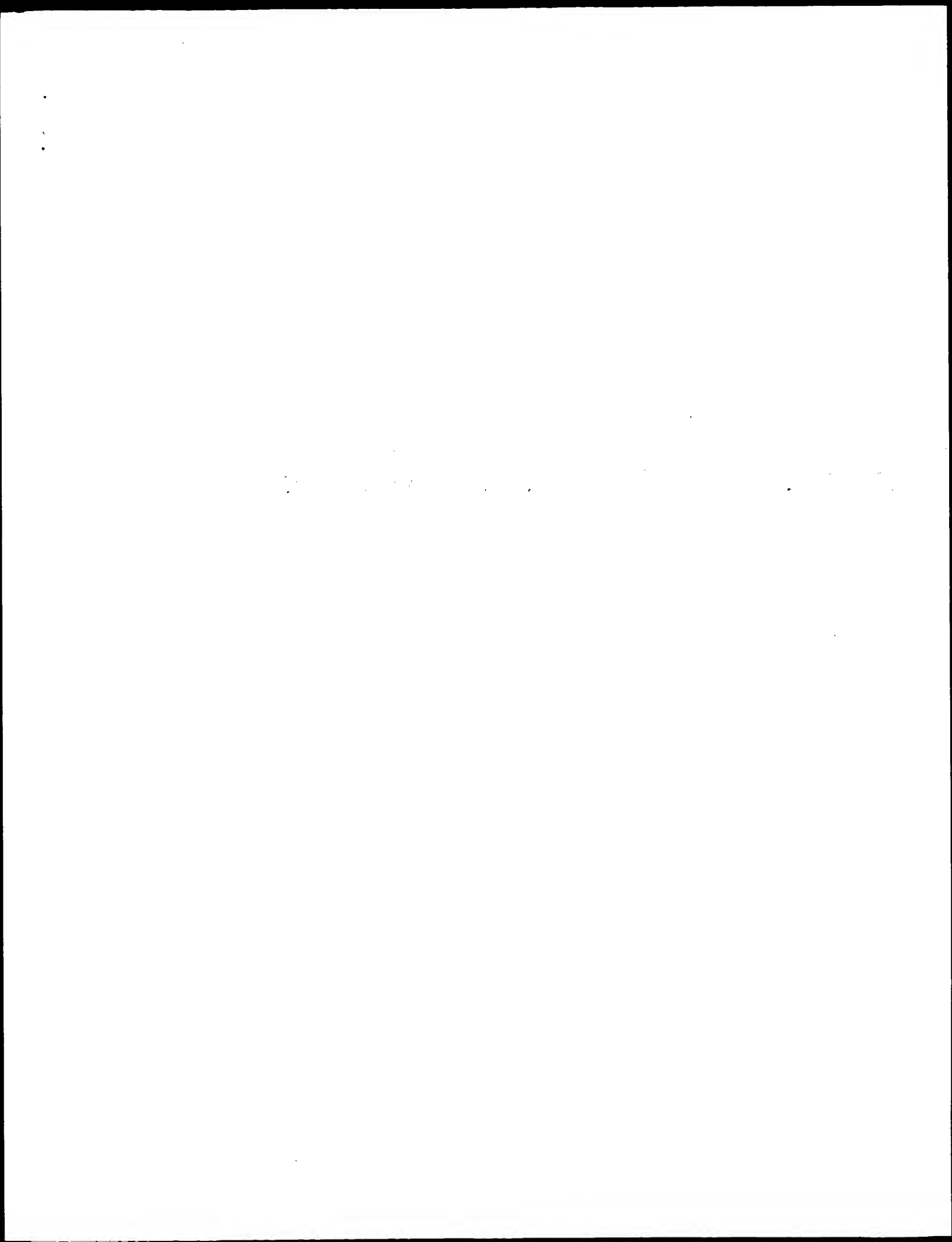
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
 CC that modulate interactions between SKP and target proteins are useful for  
 CC treating tumours.  
 SQ Sequence 569 AA:  
 Query Match 100.0%; Score 156; DB 22; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 NGHKRGIACTQYRDRLVVGSSDNTIRLW 29  
 DB 427 NGHKRGIACTQYRDRLVVGSSDNTIRLW 455  
 RESULT 15  
 AAM78583  
 ID AAM78583 standard; protein; 579 AA.  
 AC AAM78583;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human protein SEQ ID NO 1245.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 PA WO200157190-A2.  
 PN 09-AUG-2001.  
 PD  
 XX  
 PF 05-FEB-2001; 2001MO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejrtman T, Goodrich R;  
 XX WPI: 2001-476283/51.  
 DR N-PSDB; AAK51716.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 PS Claim 20; Page 3504-3505; 6221pp; English.  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78133-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX  
 SQ Sequence 579 AA;

Query Match 100.0%; Score 156; DB 22; length 579;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-16;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGHKRGIACTQYRDRLVWGSSDNTIRLW 29  
 |||  
 Db 437 NGHKRGIACTQYRDRLVWGSSDNTIRLW 465

Search completed: April 11, 2003, 11:48:16  
 Job time : 8.93666 secs





GenCore version 5.1.4-P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 3.00244 Seconds  
(without alignments)  
284.191 Million cell updates/sec

Title: US-09-601-168B-2\_COPY\_427\_455  
Perfect score: 156  
Sequence: 1 NGHKGICLQYRDLRVSSSDNTIRLM 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574 ;

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/prodata/1/1aa/6A.COMB.pep: \*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/prodata/1/1aa/PCRTUS.COMB.pep: \*  
6: /cgn2\_6/prodata/1/1aa/ackfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	156	100.0	517	1	US-08-190-802A-30
2	156	100.0	517	4	US-08-473-346-30
3	156	100.0	517	4	US-08-473-089-30
4	156	100.0	517	4	US-08-487-072A-30
5	150	96.2	23	1	US-08-190-802A-87
6	150	96.2	23	4	US-08-473-346-87
7	150	96.2	23	4	US-08-473-089-87
8	150	96.2	23	4	US-08-487-072A-87
9	87	55.8	34	1	US-08-190-802A-83
10	87	55.8	34	4	US-08-473-346-83
11	87	55.8	34	4	US-08-473-089-83
12	87	55.8	34	4	US-08-487-072A-83
13	82.5	52.9	28	1	US-08-190-802A-84
14	82.5	52.9	28	4	US-08-473-346-84
15	82.5	52.9	28	4	US-08-473-089-84
16	82.5	52.9	28	4	US-08-487-072A-84
17	78	50.0	640	4	US-09-177-165A-30
18	74.5	47.8	209	3	US-08-899-578-6
19	74.5	47.8	587	2	US-08-899-578-2
20	70	44.9	31	1	US-08-190-802A-93
21	70	44.9	31	1	US-08-190-802A-230
22	70	44.9	31	4	US-08-477-346-93
23	70	44.9	31	4	US-08-477-346-230
24	70	44.9	31	4	US-08-473-089-93
25	70	44.9	31	4	US-08-473-089-230
26	70	44.9	31	4	US-08-487-072A-93
27	70	44.9	31	4	US-08-487-072A-230

28	70	44.9	704	1	US-08-190-802A-62	Sequence 62, Appl
29	70	44.9	704	1	US-08-188-582-5	Sequence 5, Appl
30	70	44.9	704	1	US-08-646-715-5	Sequence 5, Appl
31	70	44.9	704	2	US-08-308-818-3	Sequence 3, Appl
32	70	44.9	704	4	US-08-477-346-62	Sequence 62, Appl
33	70	44.9	704	4	US-08-473-089-62	Sequence 62, Appl
34	70	44.9	704	4	US-08-487-072A-62	Sequence 31, Appl
35	70	44.9	906	1	US-08-190-802A-31	Sequence 31, Appl
36	70	44.9	906	4	US-08-477-346-31	Sequence 31, Appl
37	70	44.9	906	4	US-08-473-089-31	Sequence 31, Appl
38	70	44.9	906	4	US-08-487-072A-31	Sequence 23, Appl
39	69.5	44.6	41	4	US-08-988-856B-23	Sequence 7, Appl
40	69.5	44.6	212	3	US-08-899-578-7	Sequence 32, Appl
41	69.5	44.6	779	1	US-08-190-802A-32	Sequence 32, Appl
42	69.5	44.6	779	4	US-08-477-346-32	Sequence 32, Appl
43	69.5	44.6	779	4	US-08-473-089-32	Sequence 32, Appl
44	69.5	44.6	779	4	US-08-487-072A-32	Sequence 32, Appl
45	69.5	44.6	779	4	US-09-177-165A-29	Sequence 29, Appl

#### ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
: Sequence 30, Application US/08190802A  
: Patent No. 5519003  
GENERAL INFORMATION:  
: APPLICANT: Mochly-Rosen, Daria  
: APPLICANT: Kon, Dorit  
: TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
: TITLE OF INVENTION: Theroef  
: NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: Dehlinger & Associates  
: STREET: P.O. Box 60850  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94306-0850  
COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/190,802A  
: FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
: NAME: Fabian, Gary R.  
: REGISTRATION NUMBER: 33,875  
: REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 324-0880  
: TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 30:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 517 amino acids  
: TYPE: amino acid  
: TOPOLOGY: unknown  
: MOLECULE TYPE: peptide  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: ORIGINAL SOURCE:  
: INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-190-802A-30

Query Match 100.0% Score 156, DB 1: Length 517;  
Best Local Similarity 100.0% Pred. No. 1.3e-16;  
Matches 29: Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 NGHKGICLQYRDLRVSSSDNTIRLM 29

Db 391 NGHKGACIOLYRDRLVYSGSSDNTIRLM 419

## RESULT 2

US-08-477-346-30  
; Sequence 30, Application US/08477346  
; Patent No. 6262023  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Theof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,346  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,072  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 517 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-477-346-30

## Query Match

Best Local Similarity 100.0%; Score 156; DB 4; Length 517;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGHKGACIOLYRDRLVYSGSSDNTIRLM 29  
Db 391 NGHKGACIOLYRDRLVYSGSSDNTIRLM 419

## RESULT 3

US-08-473-089-30  
; Sequence 30, Application US/08473089  
; Patent No. 6342368  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Theof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-473-089-30

## Query Match

Best Local Similarity 100.0%; Score 156; DB 4; Length 517;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGHKGACIOLYRDRLVYSGSSDNTIRLM 29  
Db 391 NGHKGACIOLYRDRLVYSGSSDNTIRLM 419

## RESULT 4

US-08-487-072A-30  
; Sequence 30, Application US/08487072A  
; Patent No. 6423684  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Theof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,072A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-487-072A-30

Query Match 100.0%; Score 156; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 1.3e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHRKGIACIQYRDRLVYSGSSDNTIRLM 29  
DB 391 GHRKGIACIQYRDRLVYSGSSDNTIRLM 419

RESULT 5  
US-08-190-802A-87  
Sequence 87, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190.802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13  
US-08-190-802A-87

Query Match 96.2%; Score 150; DB 1; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4e-17;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHRKGIACIQYRDRLVYSGSSDNTIRLM 29

DB 1 GHRKGIACIQYRDRLVYSGSSDNTIRLM 28

RESULT 6  
US-08-477-346-87  
Sequence 87, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morlison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477.346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13  
US-08-477-346-87

Query Match 96.2%; Score 150; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4e-17;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHRKGIACIQYRDRLVYSGSSDNTIRLM 29  
DB 1 GHRKGIACIQYRDRLVYSGSSDNTIRLM 28

RESULT 7  
US-08-473-089-87  
Sequence 87, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morlison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/473.089  
APPLICATION NUMBER: US/08/473.089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13  
US-08-473-089-87

Query Match 96.2%; Score 150; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4e-17;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHKRGIACTGYRDRLVVSGSSDNTIRLM 29  
Db 1 GHKRGIACTGYRDRLVVSGSSDNTIRLM 28

RESULT 8  
US-08-487-072A-87  
Sequence 87, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theoreof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/487.072A  
APPLICATION NUMBER: US/08/487.072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP IV, Fig. 13  
US-08-487-072A-87

Query Match 96.2%; Score 150; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 4e-17;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GHKRGIACTGYRDRLVVSGSSDNTIRLM 29  
Db 1 GHKRGIACTGYRDRLVVSGSSDNTIRLM 28

RESULT 9  
US-08-190-802A-83  
Sequence 83, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theoreof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/190.802A  
APPLICATION NUMBER: US/08/190.802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0960  
TELEFAX: (415) 324-0880  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP II, Fig. 13  
US-08-190-802A-83

Query Match 55.8%; Score 87; DB 1; Length 34;  
Best Local Similarity 56.0%; Pred. No. 4.4e-07;  
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 RGIACTGYRDRLVVSGSSDNTIRLM 29

Db 9 KGVYCLQYDDQKIVSGLRDNTIKIM 33

## RESULT 10

US-08-477-346-83  
Sequence 83, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-477-346-83

Query Match 55.8%; Score 87; DB 4; Length 34;  
Best Local Similarity 56.0%; Pred. No. 4.4e-07;  
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 RGIACLOYRDLRVVSGSSDNTIRLM 29  
Db 9 KGVYCLQYDDQKIVSGLRDNTIKIM 33

## RESULT 11

US-08-473-089-83  
Sequence 83, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-473-089-83

Query Match 55.8%; Score 87; DB 4; Length 34;  
Best Local Similarity 56.0%; Pred. No. 4.4e-07;  
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 RGIACLOYRDLRVVSGSSDNTIRLM 29  
Db 9 KGVYCLQYDDQKIVSGLRDNTIKIM 33

## RESULT 12

US-08-487-072A-83  
Sequence 83, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP r1, Fig. 13  
US-08-487-072A-83

Query Match 52.8%; Score 87; DB 4; Length 34;  
Best Local Similarity 56.0%; Pred. No. 4.4e-07;  
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 5 GKIACIQYRDLVYVSSDNTIRLM 29  
DB 9 KGVYCLQYDQKIVSGLRDNTIKIW 33

RESULT 13  
US-08-190-802A-84  
Sequence 84, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP r1, Fig. 13  
US-08-190-802A-84

Query Match 52.9%; Score 82.5; DB 1; Length 28;  
Best Local Similarity 46.4%; Pred. No. 1.8e-06;  
Matches 13; Conservative 9; Mismatches 5; Indels 1; Gaps 1;

OY 2 GHKRIACIQYRDLVYVSSDNTIRLM 29

DB 1 GHKSVLCIQYDERVITIG-SDSTVRW 27

RESULT 14  
US-08-477-346-84  
Sequence 84, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP r1, Fig. 13  
US-08-477-346-84

Query Match 52.9%; Score 82.5; DB 4; Length 28;  
Best Local Similarity 46.4%; Pred. No. 1.8e-06;  
Matches 13; Conservative 9; Mismatches 5; Indels 1; Gaps 1;

OY 2 GHKRIACIQYRDLVYVSSDNTIRLM 29  
DB 1 GHKSVLCIQYDERVITIG-SDSTVRW 27

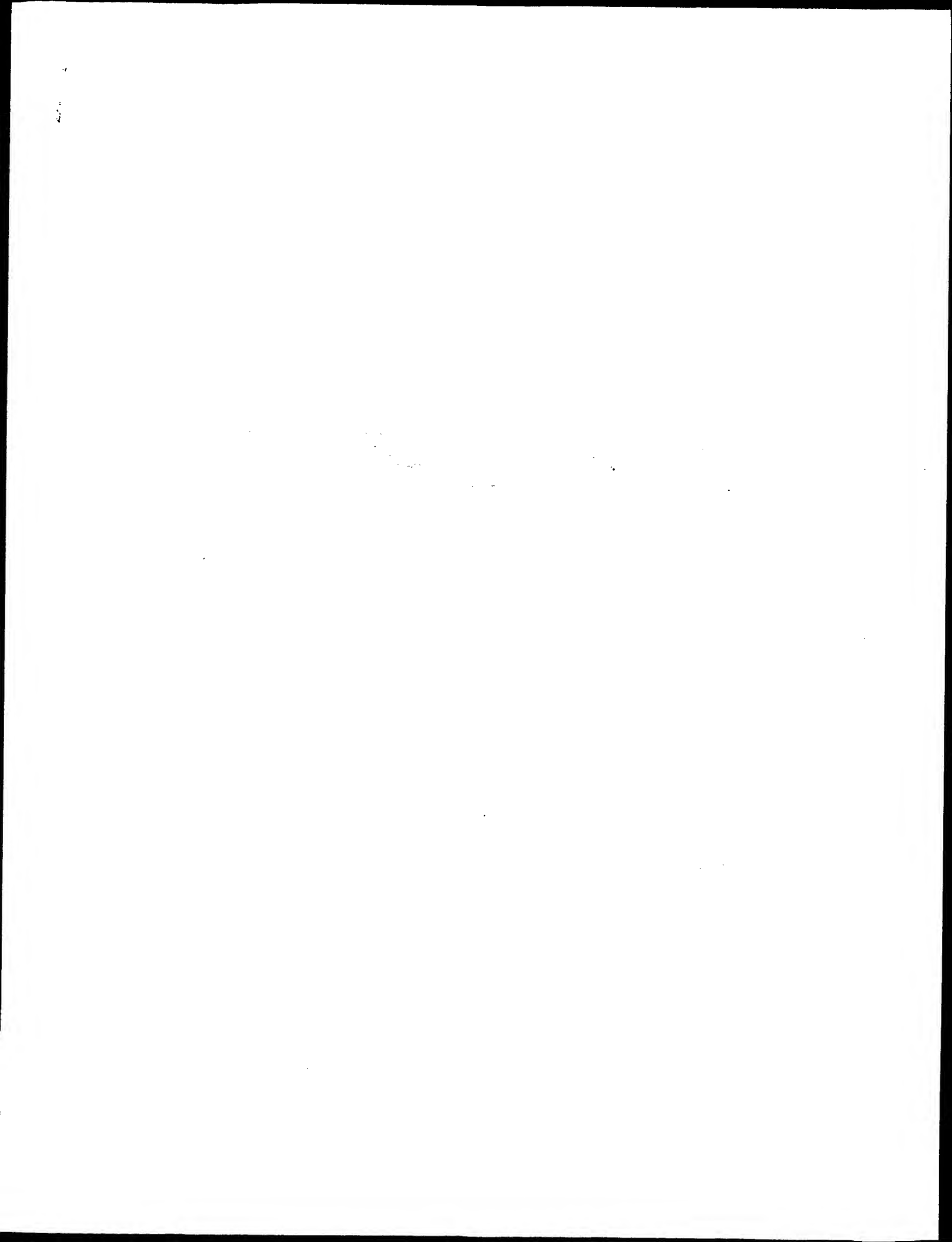
RESULT 15  
US-08-473-089-84  
Sequence 84, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP r11, Fig. 13  
US-08-473-089-84

Query Match 52.9%; Score 82.5; DB 4; Length 28;  
Best Local Similarity 46.4%; Pred. No. 1.8e-06;  
Matches 13; Conservative 9; Mismatches 5; Indels 1; Gaps 1;

OY 2 GHRGICACLOVDRRLVVGSSDNTIRLM 29  
|| : ||| :|:::| ||:|:|  
Db 1 GHTGSVLCLOYDERVITIG-SDSTVRW 27

Search completed: April 10, 2003, 12:51:19  
Job time : 3.00244 secs





GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:49:06 ; Search time 2.96711 Seconds  
(without alignments)  
597.532 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_427\_455

Perfect score: 156

Sequence: 1 NGHKGRIACLOYRDLRVVSGSSDNTIRLM 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA:\*

1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCOT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	156	100.0	569	9	US-10-038-010-8
2	156	100.0	569	12	US-10-042-417-2
3	91	58.3	678	10	US-09-801-368-314
4	79	50.6	540	9	US-09-213-888-7
5	79	50.6	540	9	US-09-213-888-10
6	79	50.6	540	9	US-09-328-877A-7
7	79	50.6	540	9	US-09-328-877A-10
8	79	50.6	545	9	US-09-213-888-6
9	79	50.6	545	9	US-09-328-877A-6
10	79	50.6	553	9	US-09-213-888-5
11	79	50.6	553	9	US-09-328-877A-5
12	79	50.6	559	9	US-09-213-888-9
13	79	50.6	559	9	US-09-328-877A-9
14	79	50.6	589	9	US-09-213-888-8
15	79	50.6	589	9	US-09-328-877A-8
16	79	50.6	592	9	US-09-213-888-4
17	79	50.6	626	9	US-09-328-877A-4
18	79	50.6	626	9	US-09-213-888-21
19	79	50.6	626	9	US-09-328-877A-21

20	79	50.6	627	9	US-09-213-888-3	Sequence 3, Appli
21	79	50.6	627	9	US-09-328-877A-3	Sequence 3, Appli
22	79	50.6	666	9	US-09-213-888-27	Sequence 27, Appli
23	79	50.6	666	9	US-09-328-877A-27	Sequence 27, Appli
24	79	50.6	669	9	US-09-213-888-25	Sequence 25, Appli
25	79	50.6	669	9	US-09-328-877A-25	Sequence 25, Appli
26	78	50.0	640	9	US-09-213-888-25	Sequence 25, Appli
27	70	44.9	744	10	US-10-060-019-30	Sequence 30, Appli
28	69.5	44.6	41	10	US-09-925-300-1347	Sequence 1347, Ap
29	69.5	44.6	41	10	US-09-843-845-23	Sequence 23, Appli
30	68.5	43.9	779	9	US-10-060-019-29	Sequence 897, App
31	68	43.6	126	9	US-10-083-357-897	Sequence 12, Appli
32	68	43.6	732	9	US-09-832-292-12	Sequence 8, Appli
33	67	42.9	1146	9	US-09-994-485-8	Sequence 10, Appli
34	67	42.9	1146	9	US-09-832-292-10	Sequence 6, Appli
35	65.5	42.0	484	9	US-10-132-744A-4	Sequence 4, Appli
36	65.5	42.0	484	9	US-10-132-744A-2	Sequence 2, Appli
37	65.5	42.0	485	9	US-10-132-744A-6	Sequence 6, Appli
38	65	41.7	44	10	US-09-843-845-26	Sequence 26, Appli
39	65	41.7	296	9	US-10-083-357-1300	Sequence 1300, Ap
40	65	41.7	713	10	US-09-843-845-20	Sequence 408, App
41	63.5	40.4	64	10	US-09-843-845-20	Sequence 36159, A
42	63	40.4	422	12	US-10-042-417-4	Sequence 4, Appli
43	63	40.4	114	9	US-09-796-692-1445	Sequence 1445, Ap
44	62	39.7	114	9	US-09-796-692-1445	Sequence 1445, Ap
45	62	39.7	114	9	US-09-796-692-1541	Sequence 1541, Ap

#### ALIGNMENTS

RESULT 1  
US-10-038-010-8 Application US/10038010  
Sequence 8, Appli  
Publication No. US20030040089A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS  
APPLICANT: Pierre, Legrain  
TITLE OF INVENTION: Protein-protein interactions in adipocyte cells  
FILE REFERENCE: B4767A  
CURRENT FILING DATE: 2002-07-23  
CURRENT FILING DATE: 2002-07-23  
PRIORITY FILING DATE: 2001-01-02  
PRIORITY FILING DATE: 2001-01-02  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: beta-TrCP1  
LOCATION: (1)..(569)  
OTHER INFORMATION:  
US-10-038-010-8  
Query Match 100.0% Score 156: DB 9: Length 569:  
Best Local Similarity 100.0% Pred. No. 2.4e-16:  
Matches 29: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
QY 1 NGHKGRIACLOYRDLRVVSGSSDNTIRLM 29  
Db 427 NGHKGRIACLOYRDLRVVSGSSDNTIRLM 455  
RESULT 2  
US-10-042-417-2  
Sequence 2, Appli Application US/10042417  
Patent No. US20020123082A1  
GENERAL INFORMATION:  
APPLICANT: Pagano, M  
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS

FILE REFERENCE: 5914-090-999  
CURRENT APPLICATION NUMBER: US/10/042.417  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-5  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-042-417-2

Query Match 100.0%; Score 156; DB 12;  
Best Local Similarity 100.0%; Pred. No. 2,4e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGRKGIACQYRDRLVYSSSDNTIRLW 29  
Db 427 NGRKGIACQYRDRLVYSSSDNTIRLW 455

RESULT 3  
US-09-801-368-314  
Sequence 314, Application US/09801368  
Patent No. US20020128250A1  
GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Call, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: Patentln version 3.0  
SEQ ID NO 314  
LENGTH: 678  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
US-09-801-368-314

Query Match 58.3%; Score 91; DB 10; Length 678;  
Best Local Similarity 53.6%; Pred. No. 4,7e-06;  
Matches 15; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 2 GHKGIACQYRDRLVYSSSDNTIRLW 29  
Db 387 GHKGIACQYRDRLVYSSSDNTIRLW 414

RESULT 4  
US-09-213-888-7  
Sequence 7, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-7

Query Match 50.6%; Score 79; DB 9; Length 540;  
Best Local Similarity 42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHKGIACQYRDRLVYSSSDNTIRLW 29  
Db 252 GHKGIACQYRDRLVYSSSDNTIRLW 279

RESULT 5  
US-09-213-888-10  
Sequence 10, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 10  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-10

Query Match 50.6%; Score 79; DB 9; Length 540;  
Best Local Similarity 42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHKGIACQYRDRLVYSSSDNTIRLW 29  
Db 252 GHKGIACQYRDRLVYSSSDNTIRLW 279

RESULT 6  
US-09-328-877A-7  
Sequence 7, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT

ORGANISM: Homo sapiens  
US-09-328-877A-7

Query Match  
Best Local Similarity 50.6%; Score 79; DB 9; Length 540;  
42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRGACIACLRDRRLVSGSSDNTIRLW 29  
DB 252 GHTGCVSSQMRDNIITISGTDRTLKW 279

RESULT 7  
US-09-328-877A-10  
Sequence 10, Application US/09328877A  
Patent No. US2002017187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-10

Query Match  
Best Local Similarity 50.6%; Score 79; DB 9; Length 540;  
42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRGACIACLRDRRLVSGSSDNTIRLW 29  
DB 252 GHTGCVSSQMRDNIITISGTDRTLKW 279

RESULT 8  
US-09-213-888-6  
Sequence 6, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 545  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-6

Query Match  
Best Local Similarity 50.6%; Score 79; DB 9; Length 545;  
42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRGACIACLRDRRLVSGSSDNTIRLW 29  
DB 257 GHTGCVSSQMRDNIITISGTDRTLKW 284

RESULT 9  
US-09-328-877A-6

Sequence 6, Application US/09328877A  
Patent No. US2002017187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 545  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-6

Query Match  
Best Local Similarity 50.6%; Score 79; DB 9; Length 545;  
42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRGACIACLRDRRLVSGSSDNTIRLW 29  
DB 257 GHTGCVSSQMRDNIITISGTDRTLKW 284

RESULT 10  
US-09-213-888-5  
Sequence 5, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-5

Query Match  
Best Local Similarity 50.6%; Score 79; DB 9; Length 553;  
42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHRGACIACLRDRRLVSGSSDNTIRLW 29  
DB 265 GHTGCVSSQMRDNIITISGTDRTLKW 292

RESULT 11  
US-09-328-877A-5  
Sequence 5, Application US/09328877A  
Patent No. US2002017187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

;; TITLE OF INVENTION: Encode Them  
;; FILE REFERENCE: 6142  
;; CURRENT APPLICATION NUMBER: US/09/328,877A  
;; CURRENT FILING DATE: 1999-06-09  
;; NUMBER OF SEQ ID NOS: 27  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 5  
;; LENGTH: 553  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-328-877A-5

Query Match 50.6%; Score 79; DB 9; Length 553;  
Best Local Similarity 42.9%; Pred. No. 0.00028;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHKRGIAQLQYRDRLVSGSSDNTIRLM 29  
Db 265 GHTGCVSSQMRDNIISGSTDRITLKYW 292

RESULT 12  
US-09-213-888-9  
; Sequence 9, Application US/09213888A  
; Patent No. US20020164683A1  
; GENERAL INFORMATION:  
; APPLICANT: Gurney, Mark E.  
; APPLICANT: Li, Jinhe  
; APPLICANT: Pauley, Adele M.  
; APPLICANT: Pharmacia & Upjohn Company  
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
; FILE REFERENCE: 6142  
; CURRENT APPLICATION NUMBER: US/09/213,888A  
; CURRENT FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 559  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-213-888-9

Query Match 50.6%; Score 79; DB 9; Length 559;  
Best Local Similarity 42.9%; Pred. No. 0.00029;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHKRGIAQLQYRDRLVSGSSDNTIRLM 29  
Db 271 GHTGCVSSQMRDNIISGSTDRITLKYW 298

RESULT 13  
US-09-328-877A-9  
; Sequence 9, Application US/09328877A  
; Patent No. US20020177187A1  
; GENERAL INFORMATION:  
; APPLICANT: Gurney, Mark E.  
; APPLICANT: Li, Jinhe  
; APPLICANT: Pauley, Adele M.  
; APPLICANT: Pharmacia & Upjohn Company  
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
; FILE REFERENCE: 6142  
; CURRENT APPLICATION NUMBER: US/09/328,877A  
; CURRENT FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 559  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-328-877A-9

Query Match 50.6%; Score 79; DB 9; Length 559;  
Best Local Similarity 42.9%; Pred. No. 0.00029;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;  
OY 2 GHKRGIAQLQYRDRLVSGSSDNTIRLM 29  
Db 271 GHTGCVSSQMRDNIISGSTDRITLKYW 298

RESULT 14  
US-09-213-888-8  
; Sequence 8, Application US/09213888A  
; Patent No. US20020164683A1  
; GENERAL INFORMATION:  
; APPLICANT: Gurney, Mark E.  
; APPLICANT: Li, Jinhe  
; APPLICANT: Pauley, Adele M.  
; APPLICANT: Pharmacia & Upjohn Company  
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
; FILE REFERENCE: 6142  
; CURRENT APPLICATION NUMBER: US/09/213,888A  
; CURRENT FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 589  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-213-888-8

Query Match 50.6%; Score 79; DB 9; Length 589;  
Best Local Similarity 42.9%; Pred. No. 0.00031;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHKRGIAQLQYRDRLVSGSSDNTIRLM 29  
Db 301 GHTGCVSSQMRDNIISGSTDRITLKYW 328

RESULT 15  
US-09-328-877A-8  
; Sequence 8, Application US/09328877A  
; Patent No. US20020177187A1  
; GENERAL INFORMATION:  
; APPLICANT: Gurney, Mark E.  
; APPLICANT: Li, Jinhe  
; APPLICANT: Pauley, Adele M.  
; APPLICANT: Pharmacia & Upjohn Company  
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
; FILE REFERENCE: 6142  
; CURRENT APPLICATION NUMBER: US/09/328,877A  
; CURRENT FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 589  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-328-877A-8

Query Match 50.6%; Score 79; DB 9; Length 589;  
Best Local Similarity 42.9%; Pred. No. 0.00031;  
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHKRGIAQLQYRDRLVSGSSDNTIRLM 29  
Db 301 GHTGCVSSQMRDNIISGSTDRITLKYW 328

Search completed: April 10, 2003, 13:16:46

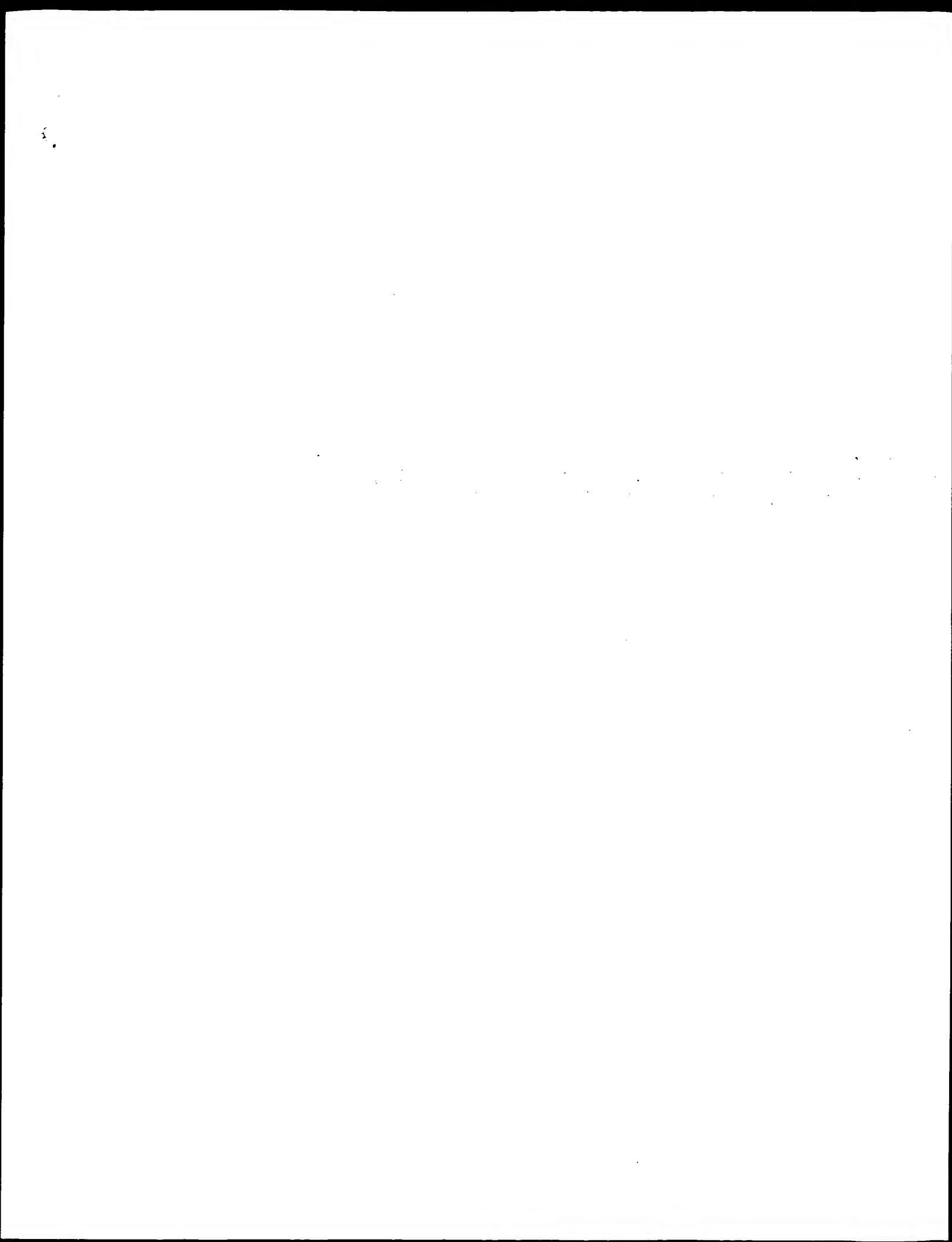
Fri Apr 11 13:31:12 2003

us-09-601-168b-2\_copy\_427\_455.rapb

Page 5

Job time : 2.96711 secs

---



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 3.35566 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_427\_455  
Perfect score: 156  
Sequence: 1 NGHKRGIAQLQYRDRLVYSGSSDNTIRLM 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	156	100.0	518 2 B48088	beta-transducin re
2	140	89.7	701 2 T16607	hypothetical prote
3	95	60.9	267 2 S62507	hypothetical trp-a
4	95	60.9	506 2 T50211	WD-repeat protein
5	87	55.8	650 2 T46660	sulfur controller-
6	86	55.1	605 2 T38932	probable sulfur me
7	79	50.6	415 2 T00642	hypothetical prote
8	79	50.6	438 2 T45823	hypothetical prote
9	78.5	50.3	358 2 T40651	pre-mrna splicing
10	78	50.0	640 2 T49932	MET30 protein - ye
11	77	49.4	651 2 T50289	WD repeat protein
12	76	48.7	465 2 T00982	En/Spm-like transp
13	76	48.1	1893 2 T22661	hypothetical prote
14	75	47.8	703 2 T43557	F-box/WD-repeat pr
15	74.5	47.8	579 2 T22703	hypothetical prote
16	73	46.8	659 2 S28108	hypothetical prote
17	73	46.8	1747 2 AC1842	WD-40 repeat prote
18	72.5	46.5	760 2 T41644	hypothetical trp-a
19	72	46.2	532 2 T18432	hypothetical prote
20	72	46.2	1201 1 ERBYA	coatmer complex a
21	70.5	45.2	446 2 S67437	Damage and replica
22	70	44.9	265 2 ARI890	WD-repeat protein
23	70	44.9	336 2 T27762	hypothetical prote
24	70	44.9	704 2 S33263	transcription init
25	70	44.9	715 2 S38051	DDAI protein - yea
26	70	44.9	906 2 S35342	Golgi-associated p
27	70	44.9	906 2 S35312	coatmer complex b
28	70	44.9	961 2 E86245	hypothetical prote
29	70	44.9	1184 2 A96638	hypothetical prote

30	69.5	44.6	779 2 S56245	cell division cont
31	69	44.2	333 2 G85034	probable WD-repeat
32	69	44.2	777 2 T41075	hypothetical WD-re
33	68.5	43.9	473 2 T33805	hypothetical prote
34	68.5	43.9	601 2 T38258	coronin-like prote
35	68	43.6	495 2 T04784	hypothetical prote
36	68	43.6	614 2 S58306	WD-40 repeat regul
37	68	43.6	755 2 T00066	hypothetical prote
38	67	42.9	466 2 T45585	hypothetical prote
39	67	42.9	1146 2 A55532	myosin-heavy-chain
40	67	42.9	1224 1 ERH04H	coatmer complex a
41	67	42.9	1711 2 AD1842	WD-40 repeat prote
42	66.5	42.6	515 2 S19487	hypothetical prote
43	66	42.3	559 2 AB2202	hypothetical prote
44	66	42.3	T21970	hypothetical prote
45	65.5	42.0	532 2 F84901	probable photomorp

## ALIGNMENTS

RESULT 1  
B48088  
beta-transducin repeat-containing protein - African clawed frog  
N:Alternate names: beta-Trp  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000  
C:Accession: B48088  
R:Spevak, W.; Keiper, B.D.; Stratawa, C.; Castanon, M.J.  
Mol. Cell. Biol. 13, 4953-4966, 1993  
A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase  
A:Reference number: A48088; MUID:93330289; PMID:8393141  
A:Accession: B48088  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-518 <SP>  
A:Cross-references: GR:98268; NID:9295542; PIDN:AAA02810.1; PID:9295543  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: duplication  
F:431-462/Domain: WD repeat homology <WD1>

Query Match 100.0%; Score 156; DB 2; Length 518;  
Best Local Similarity 100.0%; Pred. No. 3.3e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 392 NGHKRGIAQLQYRDRLVYSGSSDNTIRLM 420  
OY 1 NGHKRGIAQLQYRDRLVYSGSSDNTIRLM 29  
|||||

RESULT 2  
T16607  
hypothetical protein K10B2.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
R:Miller, N.  
submitted to the EMBL Data Library, June 1995  
A:Description: The sequence of C. elegans cosmid K10B2.  
A:Reference number: Z18545  
A:Accession: T16607  
A>Status: preliminary; translated from GR/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-701 <ML>  
A:Cross-references: EMBL:028730; NID:9860694; PID:9860695; PIDN:AAA68258.1; CESP:K10B  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:K10B2.1  
A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 89.7%; Score 140; DB 2; Length 701;  
Best Local Similarity 92.9%; Pred. No. 1.4e-13;  
Matches 26; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GHKRGIACTQYRDRLVVGSSDNTIRLM 29  
 ||:||||||| |||||||||||||  
 Db 419 GHKRGIACTQYRDRLVVGSSDNTIRLM 446

## RESULT 3

S62507

hypothetical trp-asp repeat-containing protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000

C:Accession: T38502; S62507  
 R:Jones, L.; Murphy, L.; McNeill, A.; Simpson, I.; Harris, D.; Barrell, B.G.; Rajandream,

submitted to the EMBL Data Library, October 1995

A:Reference number: Z21798

A:Accession: T38502

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1267 <02>

A:Cross-references: EMBL:Z66525; NID:g1044926; PIDN:CAA91423.1; PID:g1044927; GSFDB:GN00

C:Genetics:

A:Map position: 1

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

F:18-51/Domain: WD repeat homology <WD1>

F:104-137/Domain: WD repeat homology <WD2>

## Query Match

Best Local Similarity 60.9%; Score 95; DB 2; Length 267;  
 Matches 17; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 GHKRGIACTQYRDRLVVGSSDNTIRLM 29  
 ||:||||||| |||||||||||||

Db 147 HAHORGIACQYNGKRTIVSGSSDLTIRLF 175

## RESULT 4

T50211

WD-repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
 C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000

C:Accession: T50211  
 R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.; Mc

submitted to the EMBL Data Library, January 2000

A:Reference number: Z25046

A:Accession: T50211

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1506 <MCD>

A:Cross-references: EMBL:AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05

A:Experimental source: strain 972h(-); cosmid G30

C:Genetics:

A:Gene: SPAC29E6.01; SPDB:SPAC30.05

A:Map position: 1

A:Introns: 43/1; 74/3

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

## Query Match

Best Local Similarity 60.9%; Score 95; DB 2; Length 506;  
 Matches 17; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 GHKRGIACTQYRDRLVVGSSDNTIRLM 29  
 ||:||||||| |||||||||||||

Db 386 HAHORGIACQYNGKRTIVSGSSDLTIRLF 414

## RESULT 5

T46660

sulfur controller-2 protein [imported] - Neurospora crassa

C:Species: Neurospora crassa  
 C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000

R:Kumar, A.; Paietta, J.V.

Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995

A:Title: The sulfur controller-2 negative regulatory gene of Neurospora crassa encode  
 A:Reference number: Z23121; MIDID:95241499; PMID:7724564  
 A:Accession: T46660

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1650 <KUM>

A:Cross-references: EMBL:U17251; NID:g806757; PIDN:AAA68968.1; PID:g806758

C:Genetics:

A:Gene: scon-2

A:Map position: 3

A:Introns: 75/3; 319/1; 354/1

C:Function:

A:Description: negatively regulates sulfur structural gene expression

A>Note: scon-2+ expression is dependent on CYS3 function and the binding of CYS3 to t

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

## Query Match

Best Local Similarity 55.8%; Score 87; DB 2; Length 650;  
 Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 GHKRGIACTQYRDRLVVGSSDNTIRLM 29  
 ||:||||||| |||||||||||||

Db 292 GHKRGIACTQYRDRLVVGSSDNTIRLM 319

## RESULT 6

T38932

probable sulfur metabolite control protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 26-May-2000

C:Accession: T38932  
 R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z21818

A:Accession: T38932

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1605 <BAD>

A:Cross-references: EMBL:Z94864; PIDN:CAB08168.1; GSPDB:GN00066; SPDB:SPAC57A10.05C

A:Experimental source: strain 972h(-); cosmid c57A10

C:Genetics:

A:Gene: SPDB:SPAC57A10.05C

A:Map position: 1

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

## Query Match

Best Local Similarity 55.1%; Score 86; DB 2; Length 605;  
 Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 GHKRGIACTQYRDRLVVGSSDNTIRLM 29  
 ||:||||||| |||||||||||||

Db 311 GHKRGIACTQYRDRLVVGSSDNTIRLM 338

## RESULT 7

T00642

hypothetical protein F316.5 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 22-Oct-1999

C:Accession: T00642  
 R:Federpiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Arau

submitted to the EMBL Data Library, February 1998

A:Reference number: Z14197

A:Accession: T00642

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1415 <FED>

A:Cross-references: EMBL:AC002396; NID:g2749918; PID:g2829890; GSPDB:GN00059; ATSP:F3

C:Genetics:

A:Gene: ATSP:F316.5

A:Map position: 1





A:Accession: T00982  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-465 <R0U>  
 A:Cross-references: EMBL:AC002505; NID:g2739359; PID:g2739374  
 A:Experimental source: cultivar Columbia  
 R:Rillu, X.; Kaul, S.; Rounsley, S.D.; Shee, T.F.; Bentto, M.L.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 euss, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: B84661  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-465 <STO>  
 A:Cross-references: GB:AE002093; NID:g2739374; PIDN:AA01498.1; GSPDB:GND0139  
 A:Genetics:  
 A:Gene: At2g26490; T9J22.16  
 A:Map position: 2

Query Match 48.7%; Score 76; DB 2; Length 465;  
 Best Local Similarity 50.0%; Pred. No. 0.00093;  
 Matches 14; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 2 GHRGACIACLODRRLVVGSSSDNTIRLM 29  
 ||| : ||: ||||| : ||: ||  
 Db 336 GHKLAVLCLEVAGSLVFGSADKTCICW 363

RESULT 13  
 T22661  
 hypothetical protein T01G1.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T22661; T24305  
 R:White, S.; Mortimore, B.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19594  
 A:Accession: T22661  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1893 <WIL>  
 A:Cross-references: EMBL:Z82271; PIDN:CA05214.1; GSPDB:GND0022; CESP:T01G1.1  
 A:Experimental source: clone F54E12  
 R:Kershaw, J.  
 submitted to the EMBL Data Library, March 1997  
 A:Reference number: Z19872  
 A:Accession: T24305  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1893 <W12>  
 A:Cross-references: EMBL:Z92811; PIDN:CA07273.1; GSPDB:GND0022; CESP:T01G1.1  
 A:Experimental source: clone T01G1  
 C:Genetics:  
 A:Gene: CESP:T01G1.1  
 A:Map position: 4  
 A:introns: 65/3; 101/3; 194/3; 269/3; 299/2; 373/3; 481/3; 576/3; 615/2; 704/3; 839/1;

Query Match 48.7%; Score 76; DB 2; Length 1893;  
 Best Local Similarity 42.9%; Pred. No. 0.0042;  
 Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 2 GHRGACIACLODRRLVVGSSSDNTIRLM 29  
 ||| : ||: ||||| : ||: ||  
 Db 1570 GHARGVLSVDVNEKLVGSKORTAKLM 1597

RESULT 14  
 T43557  
 F-box/WD-repeat protein pop2 - fission yeast (*Schizosaccharomyces pombe*)  
 N:Alternate names: proteolysis factor stud1  
 C:Species: *Schizosaccharomyces pombe*

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
 C:Accession: T43557; T38794; T43798  
 R:Wolff, D.A.; Jackson, P.K.  
 submitted to the EMBL Data Library, December 1997  
 A:Description: Fission yeast pop2 encodes a novel F-box/WD-repeat protein involved in  
 A:Reference number: Z22576  
 A:Accession: T43557  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-703 <MOL>  
 A:Cross-references: EMBL:AF038867; PIDN:AA095480.1  
 A:Experimental source: strain h-972  
 R:Gentles, S.; Churcher, C.M.; Bartell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z21812  
 A:Accession: T38794  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-703 <GEN>

A:Cross-references: EMBL:Z98602; PIDN:CA011275.1; GSPDB:GND0066; SPDB:SPAC4D7.03  
 A:Experimental source: strain 972h; cosmid C4D7  
 R:Jalilepalli, P.V.; Tien, D.; Kelly, T.J.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 8159-8164, 1998  
 A:Title: Stud1 targets cyclin-dependent kinase-phosphorylated Cdc18 and Rnm1 proteins  
 A:Reference number: Z22686; MUID:98318628; PMID:9653157  
 A:Accession: T43798  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-703 <JAL>  
 A:Cross-references: EMBL:AF064515; NID:g3293382; PIDN:AA039496.1; PID:g3293383  
 C:Genetics:  
 A:Gene: pop2; stud1; SPAC4D7.03  
 A:Map position: 1  
 C:Function:  
 A:Description: required to prevent spontaneous re-replication

Query Match 48.1%; Score 75; DB 2; Length 703;  
 Best Local Similarity 42.9%; Pred. No. 0.0021;  
 Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 2 GHRGACIACLODRRLVVGSSSDNTIRLM 29  
 ||| : ||: ||||| : ||: ||  
 Db 389 GHRGAVAKVAKIHENLVGSDIKTVRW 416

RESULT 15  
 T22703  
 hypothetical protein F55B12.3 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T22703  
 R:Sims, M.  
 submitted to the EMBL Data Library, September 1996  
 A:Reference number: Z19602  
 A:Accession: T22703  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-579 <WIL>  
 A:Cross-references: EMBL:Z79757; PIDN:CA02129.1; GSPDB:GND0023; CESP:F55B12.3  
 A:Experimental source: clone F55B12  
 C:Genetics:  
 A:Gene: CESP:F55B12.3  
 A:Map position: 5  
 A:introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

Query Match 47.8%; Score 74.5; DB 2; Length 579;  
 Best Local Similarity 44.8%; Pred. No. 0.002;  
 Matches 13; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

OY 2 GHR-RTGACIACLODRRLVVGSSSDNTIRLM 29  
 ||| : ||: ||||| : ||: ||  
 Db 245 GHEDHVTICMQIHDDVLTGSDNTLRKW 273

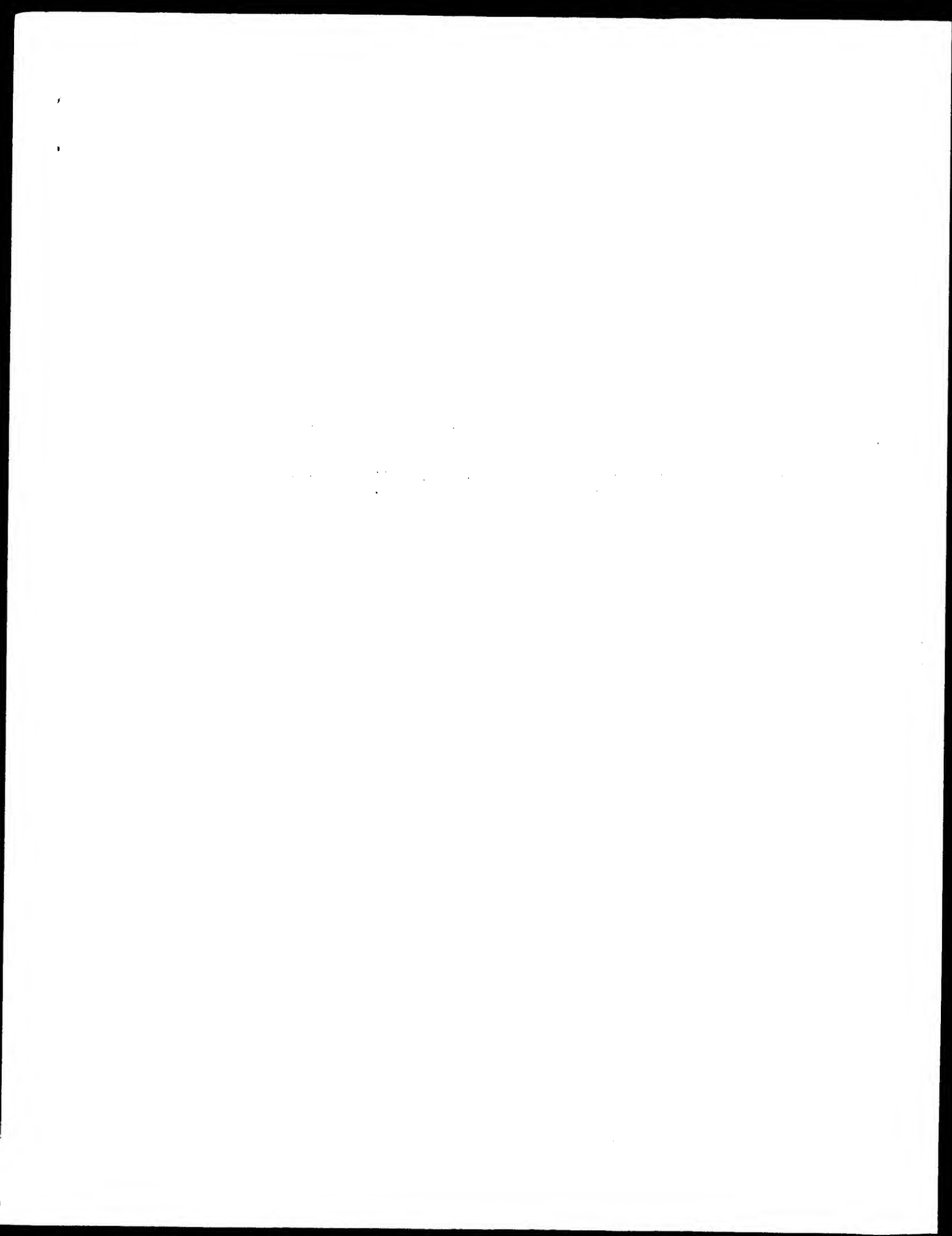
Fri Apr 11 13:31:12 2003

us-09-601-168b-2\_copy\_427\_455.rpr

Page 5

Search completed: April 10, 2003, 13:23:04  
Job time : 4.35566 secs

---



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:51:26 ; Search time 2.08404 Seconds  
(without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_427\_455

Sequence: 1 NGHKRGICLQYRDLVWVGSSDNTIRLM 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	156	100.0	518	TRCB_XENLA	Q91854 xenopus lae
2	156	100.0	542	FW1B_HUMAN	Q9UKX1 homo sapien
3	156	100.0	605	FW1A_HUMAN	Q9Y297 homo sapien
4	140	89.7	665	LI23_CAEEL	Q09990 caenorhabd
5	95	60.9	506	POEB_SCHPO	Q09855 schizosacch
6	91	58.3	678	SCOB_MENM	Q00659 emericella
7	87	55.8	650	SC02_MENM	Q01277 neurospora
8	86	55.1	640	POE1_SCHPO	P97053 schizosacch
9	78	50.0	640	MT30_YEAST	Q9P713 schizosacch
10	77	49.4	651	Y1X1_SCHPO	P53699 candida alb
11	76	48.7	684	CC4_CANAL	Q04170 schizosacch
12	75	48.1	703	SE10_CAEEL	Q93794 caenorhabd
13	74.5	47.8	579	YK16_YEAST	P36130 saccharomyc
14	73	46.8	659	PRP5_SCHPO	Q13615 schizosacch
15	72	46.2	473	PRP5_SCHPO	P53699 candida alb
16	72	46.2	1201	COPA_YEAST	Q9Y318 drosophila
17	71	45.5	361	MDS_DROME	Q10272 schizosacch
18	70.5	45.2	446	CRB3_SCHPO	P49846 drosophila
19	70	44.9	704	T2D4_DROME	P36037 saccharomyc
20	70	44.9	715	DOA1_YEAST	P36605 bos taurus
21	70	44.9	905	COPB_BOVIN	P36605 bos taurus
22	70	44.9	905	COPB_HUMAN	O55029 mus musculu
23	70	44.9	905	COPB_MOUSE	P07834 saccharomyc
24	69.5	44.6	779	CC4_YEAST	P78706 neurospora
25	69	44.2	604	RCO1_NEUCR	O35142 rattus norv
26	69	44.2	904	COPB_RAT	O13923 schizosacch
27	68.5	43.9	601	CORP_SCHPO	Q23189 hydra atten
28	68	43.6	317	GBLP_HYDAT	O099715 schizosacch
29	68	43.6	614	TU11_SCHPO	P90648 dictyostell
30	68	43.6	732	KMH8_DICDI	P42527 dictyostell
31	67	42.9	1146	KMHA_BOVIN	O27954 bos taurus
32	67	42.9	1224	COPA_BOVIN	P53621 homo sapien
33	67	42.9	1224	COPA_HUMAN	

34	66.5	42.6	515	1	YCW2_YEAST	P25382 saccharomyc
35	66	42.3	1000	1	COPB_CAEEL	Q20168 caenorhabd
36	65	41.7	432	1	WDPI_HUMAN	Q9BV38 homo sapien
37	65	41.7	713	1	TUPL_YEAST	P16649 saccharomyc
38	65	41.7	800	1	T2D4_HUMAN	Q15542 homo sapien
39	64.5	41.3	578	1	YDEM_CAEEL	Q19124 caenorhabd
40	64	41.0	334	1	WDR5_HUMAN	Q9UGP9 homo sapien
41	64	41.0	376	1	YKY4_CAEEL	Q17963 caenorhabd
42	64	41.0	535	1	SIF2_YEAST	P38262 saccharomyc
43	64	41.0	714	1	YUJ2_YEAST	P47025 saccharomyc
44	64	41.0	910	1	PWP2_CAEEL	P91341 caenorhabd
45	64	41.0	1258	1	YS00_ANASP	O8YTC2 anabaena sp

## ALIGNMENTS

```

RESULT 1
ID TRCB_XENLA STANDARD; PRT; 518 AA.
AC Q91854; P70037; P70038;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-TTCP (Beta-transducin repeat-containing protein).
GN FBXW1 OR BTCP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipodae; Pipidae;
OC Xenopodinae; Xenopus.
ON NCBI_TaxID=8335;
RX MEDLINE=93330289; PubMed=8393141;
RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
RT anaphase are rescued by Xenopus CDNs encoding N-ras or a protein
RT with beta-transducin repeats."
RL Mol. Cell. Biol. 13:4953-4966(1993).
RN [2]
RP SEQUENCE OF 302-518 FROM N.A.
RX MEDLINE=97109804; PubMed=8952061;
RA Hudson J.W., Alarcon V.B., Ellson R.P.;
RT "Identification of new localized RNAs in the Xenopus oocyte by
RT differential display PCR."
RL Dev. Genet. 19:190-198(1996).
CC - FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC - SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC - DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
CC TADPOLE EMBRYO.
CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC - SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M98268; AAA02810.1; -
CC EMBL: U63921; AAB49671.1; -
CC EMBL: U63922; AAB49672.1; -
CC Interpro: IPR001810; F-box.
CC Interpro: IPR001680; WD40.
CC Pfam: PF00400; WD40; 7.

```



RA Yaron A., Hatzubal A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.:  
 RT "Identification of the receptor component of the IkappaBalpha-  
 RT ubiquitin ligase." Nature 396:590-594(1998).  
 RL Nature 396:590-594(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Lymphoid;  
 RX MEDLINE=98325370; PubMed=9660940;  
 RA Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,  
 RA Thomas D., Strehel K., Benarous R.:  
 RT "A novel human WD protein, h-beta Trcp, that interacts with HIV-1 Vpu  
 RT connects CD4 to the ER degradation pathway through an F-box motif." Mol.  
 RL Cell 11:565-574(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Cenciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.:  
 RT "Identification of a family of human F-box proteins." Curr. Biol. 9:1177-1179(1999).  
 RL [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99145464; PubMed=9990852;  
 RA Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,  
 RA Harper J.W.:  
 RT "The SCF(beta-Trcp)-ubiquitin ligase complex associates specifically  
 RT with phosphorylated destruction motifs in I-kappa-B-alpha and  
 RT beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro." Genes  
 RL Dev. 13:270-283(1999).  
 CC -1- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA  
 CC (UBIQUITINATION AND DEGRADATION).  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2: ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF101784; AAD08702.1; -  
 CC EMBL: Y14153; CA74572.1; -  
 CC EMBL: AF129530; AAF04464.1; -  
 CC Genew: HGNC:1144; BTRC.  
 CC MIM: 603482; -  
 CC InterPro: IPR001810; F-box.  
 CC InterPro: IPR001680; WD40.  
 CC Pfam: PF00440; WD40; 7.  
 CC PRINTS: PR00320; GPROTEINRPT.  
 CC PRODOM: PD000018; WD40; 4.  
 CC SMART: SM00256; FBOX; 1.  
 CC SMART: SM00320; WD40; 7.  
 CC PROSITE: PSS0181; FBOX; 1.  
 CC PROSITE: PSS0678; WD\_REPEATS\_1; 6.  
 CC PROSITE: PSS0082; WD\_REPEATS\_2; 7.  
 CC PROSITE: PSS0294; WD\_REPEATS\_REGION; 1.  
 CC Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 KW DOMAIN 190 228 F-BOX.  
 FT REPEAT 301 338 WD 1.  
 FT REPEAT 341 378 WD 2.  
 FT REPEAT 381 418 WD 3.  
 FT REPEAT 424 461 WD 4.  
 FT REPEAT 464 503 WD 5.  
 FT REPEAT 505 541 WD 6.

FT REPEAT 553 590 WD 7.  
 FT VARSPIC 17 52 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 605 AA: 6886 MW: 467437E400ED37 CRC64:  
 Query Match 100.0%; Score 156; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-17;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NGHKGACIACIYRDRLVSSGSDNTIRLM 29  
 Db 463 NGHKGACIACIYRDRLVSSGSDNTIRLM 491  
 RESULT 4  
 LI23\_CAEEL STANDARD; PRT; 665 AA.  
 ID LI23\_CAEEL STANDARD; PRT; 665 AA.  
 AC Q09990; Q8GNN6;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein lin-23.  
 GN LIN-23 OR K10B2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, DEVELOPMENTAL STAGE, AND MUTAGENESIS.  
 RX MEDLINE=20515608; PubMed=11060233;  
 RA Kipreos E.T., Gohel S.P., Hedgecock E.M.:  
 RT "The Caenorhabditis elegans F-box/WD-repeat protein lin-23 functions  
 RT to limit cell division during development." Development 127:5071-5082(2000).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Miller N.:  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.:  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Functions cell autonomously to negatively regulate cell  
 CC cycle progression. Required to restrain cell proliferation in  
 CC response to developmental cues. Probably recognizes and binds to  
 CC some proteins and promotes their ubiquitination and degradation  
 CC (by similarity).  
 CC -1- SUBUNIT: Part of a SCF (SKP1-CULLIN-F-BOX) protein ligase complex  
 CC (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).  
 CC -1- DEVELOPMENTAL STAGE: Highest levels in embryos and adults, lowest  
 CC levels in larvae. Maternal expression results in high zygotic  
 CC levels.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO X.LAEVIS FBXW1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF275253; AAG28037.1; -  
 CC EMBL: U28730; AAA68258.2; -  
 CC Wormpep: K10B2.1; CE28600.  
 CC InterPro: IPR001810; F-box.  
 CC InterPro: IPR001680; WD40.  
 CC Pfam: PF00440; WD40; 7.  
 CC PRINTS: PR00320; GPROTEINRPT.

DR PRODOM: PD000018; WD40: 4.  
 DR SMART: SM00256; FBOX: 1.  
 DR SMART: SM00320; WD40: 7.  
 DR PROSITE: PS50181; FBOX: 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1: 5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2: 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION: 1.  
 DR Developmental protein; Cell cycle; Cell division;  
 KW Ub1 conjugation pathway; Repeat; WD repeat.  
 FT DOMAIN 81 127  
 FT REPEAT 220 257 WD 1.  
 FT REPEAT 260 299 WD 2.  
 FT REPEAT 301 337 WD 3.  
 FT REPEAT 343 380 WD 4.  
 FT REPEAT 383 420 WD 5.  
 FT REPEAT 423 460 WD 6.  
 FT REPEAT 472 509 WD 7.  
 FT DOMAIN 525 582  
 FT MUTAGEN 441 441  
 SQ SEQUENCE 665 AA; 75916 MW; BFB9P9AF51P12ECC CRC64;  
 G->R: IN LIN-32(RH293).  
 Query Match 89.7%; Score 140; DB 1; Length 665;  
 Best Local Similarity 92.9%; Pred. No. 3e-14; Mismatches 1; Indels 0; Gaps 0;  
 Matches 26; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 GHRGIACTQYRRLVYSSSDNTIRLW 29  
 DB 383 GHRGIACTQYRRLVYSSSDNTIRLW 410  
 ID POFB\_SCHPO STANDARD; PRT; 506 AA.  
 AC 009855; 09P7V1;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein pof1.  
 GN POF1 OR SPAC29E6.01 OR SPAC30.05.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harrison C.L., Toda T.;  
 RT "Systematic genome-wide analysis of F-box protein-encoding genes in  
 RL fission yeast.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RX MEDLINE-21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Wellens I., Vanslechts E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Carrutti L., Lowe T., McComble W.R., Paulsen I., Potashkin J.,  
 RA Shipkavski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD\_REPEATS (TRP-ASP DOMAINS).  
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AB061694; BAB5543.1; -;  
 DR EMBL: AL136538; CAB66464.1; -;  
 DR EMBL: 265525; CA91423.1; -;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40: 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PRO0320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40: 3.  
 DR SMART: SM00256; FBOX: 1.  
 DR SMART: SM00320; WD40: 8.  
 DR PROSITE: PS50181; FBOX: 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1: 3.  
 DR PROSITE: PS50082; WD\_REPEATS\_2: 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION: 1.  
 KW Repeat; WD repeat.  
 FT DOMAIN 70 116  
 FT REPEAT 219 256 WD 1.  
 FT REPEAT 259 298 WD 2.  
 FT REPEAT 301 338 WD 3.  
 FT REPEAT 345 386 WD 4.  
 FT REPEAT 388 426 WD 5.  
 FT REPEAT 427 464 WD 6.  
 FT REPEAT 468 505 WD 7.  
 SQ SEQUENCE 506 AA; 58257 MW; CEF34DAEFEC2E10 CRC64;  
 Query Match 60.9%; Score 95; DB 1; Length 506;  
 Best Local Similarity 58.6%; Pred. No. 2.9e-07;  
 Matches 17; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 NGHRGIACTQYRRLVYSSSDNTIRLW 29  
 DB 386 HAHORGIACTQYRRLVYSSSDNTIRLW 414  
 ID SCOB\_EMENT STANDARD; PRT; 678 AA.  
 AC 000659;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sulfur metabolite repression control protein.  
 GN SCOB OR MABP1.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Natorff R.;  
 RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE  
 CC REPRESSION.  
 CC -1- SIMILARITY: CONTAINS 8 WD\_REPEATS (TRP-ASP DOMAINS).  
 CC -----



```

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U21220; AAC15905.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF004400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Transcription regulation; Repeat; WD repeat.
KW DOMAIN
FT REPEAT 178 224 F-BOX.
FT REPEAT 347 375 WD 1.
FT REPEAT 387 415 WD 2.
FT REPEAT 427 455 WD 3.
FT REPEAT 466 496 WD 4.
FT REPEAT 508 543 WD 5.
FT REPEAT 553 595 WD 6.
FT REPEAT 607 635 WD 7.
FT REPEAT 647 675 WD 8.
SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;

Query Match 58.3%; Score 91; DB 1; Length 678;
Best local similarity 53.6%; Pred. No. 1.7e-06;
Matches 15; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 GHKRGIAQLQYRDRLVSSGSSDNTIRLW 29
DB 387 GHESGIRCLQFDFTKLSSMDRTIKW 414

RESULT 7
SC02_NEUCR STANDARD; PRT; 650 AA.
AC 001277;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sulfur controller-2 (SCON2).
GN SCON-2.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RX MEDLINE=95241459; PubMed=7724564;
RA Kumar A., Patel J.V.;
RA "The sulfur controller-2 negative regulatory gene of Neurospora
RA crassa encodes a protein with beta-transducin repeats.";
RA Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
RL -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC -1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U17251; AAA68968.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF004400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Transcription regulation; Repeat; WD repeat.
KW DOMAIN
FT REPEAT 124 170 F-BOX.
FT REPEAT 292 320 WD 1.
FT REPEAT 332 360 WD 2.
FT REPEAT 372 400 WD 3.
FT REPEAT 411 441 WD 4.
FT REPEAT 453 488 WD 5.
FT REPEAT 528 564 WD 6.
FT REPEAT 576 604 WD 7.
FT REPEAT 616 644 WD 8.
SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 55.8%; Score 87; DB 1; Length 650;
Best local similarity 46.4%; Pred. No. 7.1e-06;
Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 GHKRGIAQLQYRDRLVSSGSSDNTIRLW 29
DB 292 GHENGVTCQLDDNLTATSGSYDTTIKW 319

RESULT 8
POF1_SCHPO STANDARD; PRT; 605 AA.
ID POF1_SCHPO
AC P87053;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof1 (Skp1-binding protein 1).
GN POF1 OR SBP1 OR SPAC57A10.05C.
OS Schizosaccharomyces pombe (fission yeast).
OC Schizosaccharomycetes; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Katayama S., Kitamura K., Toda T.;
RA "Systematic genome-wide analysis of F-box protein-encoding genes in
RA fission yeast.";
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

```

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Moorey P., Mole S., Mingall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares S., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Filiz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Petashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,  
 RA "the genome sequence of Schizosaccharomyces pombe";  
 RL Nature 415:671-880(2002).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB032410; BAA84528.1; -  
 DR EMBL: Z94864; CAB08168.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 5.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PS00678; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_3; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW UniProt conjugation; Repeat; WD repeat.  
 FT DOMAIN 107 153  
 FT REPEAT 271 299  
 FT REPEAT 311 339  
 FT REPEAT 350 379  
 FT REPEAT 390 420  
 FT REPEAT 432 460  
 FT REPEAT 472 500  
 FT REPEAT 510 538  
 FT REPEAT 510 538  
 SO SEQUENCE 605 AA; 67110 MW; 711809379BEC5C1F0 CRC64;

Query Match 55.18; Score 86; DB 1; Length 605;  
 Best Local Similarity 50.04; Pred. No. 9.4e-06;  
 Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 GHKRGIAQYRDRIYVSGSSDNTIRLM 29  
 Db 311 GHSSGVTCIAPDCKLISGSMKTRIM 338

RESULT 9  
 MT30\_YEAST STANDARD; PRT; 640 AA.  
 ID MT30\_YEAST

AC P39014;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE MET30 protein.  
 GN MET30 OR Y11046W.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-X2180-1A;  
 RX MEDLINE=96069360; PubMed=8524217;  
 RA Thomas D., Kuras L., Barbey R., Charest H., Blaiseau P.L.,  
 RA Surdin-Kerjan Y.,  
 RT "Met30p, a yeast transcriptional inhibitor that responds to S-  
 RT adenosylmethionine, is an essential protein with WD40 repeats";  
 RL Mol. Cell. Biol. 15:6526-6534(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288c / AB972;  
 RA Barrett B.G., Badcock K., Bankier A.T., Boyman S., Brown D.,  
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Hornsby T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis E., Lye G., Mole S., Moulton T., Odell C., Pearson D.,  
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS  
 CC GENES EXPRESSION.  
 CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT  
 CC PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z46861; CAA86905.1; -  
 DR EMBL: L26505; AAA96717.1; -  
 DR SGD: S0001308; MET30.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS00678; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS00678; WD\_REPEATS\_3; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Cysteine biosynthesis; Repeat; WD repeat.  
 FT DOMAIN 181 227  
 FT REPEAT 300 328  
 FT REPEAT 340 368  
 FT REPEAT 380 408  
 FT REPEAT 419 449  
 FT REPEAT 461 499  
 FT REPEAT 509 538  
 FT REPEAT 550 578  
 FT REPEAT 607 635  
 FT REPEAT 61 61  
 SO SEQUENCE 640 AA; 72835 MW; 5135DABCA2E1B97 CRC64;



```

FT REPEAT 478 506 WD 5.
FT REPEAT 519 549 WD 6.
FT REPEAT 561 589 WD 7.
SQ SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;

Query Match
Best local Similarity 42.9%; Pred. No. 0.00041;
Matches 12; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 2 GHKRGIAQYRDRLVVGSSDNTIRLW 29
DQ 363 GHEGVAWKXHTGNTLVGTSTDRVRW 390

RESULT 12
POP2_SCHPO STANDARD; PRT; 703 AA.
ID POP2_SCHPO
AC 014170;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE WD-repeat protein pop2 (Proteolysis factor stud1).
GN POP2 OR SUD1 OR SPAC4D7.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=99227353; PubMed=10209119;
RA Wolf D.A., McKeon F., Jackson P.K.;
RT "F-box/WD-repeat proteins pop1 and Stud1/Pop2p form complexes that
RT bind and direct the proteolysis of cdc18p."
RL Curr. Biol. 9:373-376(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=98318628; PubMed=9653157;
RA Jallepalli P.V., Tien D., Kelly T.J.;
RT "Stud1 targets cyclin-dependent kinase-phosphorylated Cdc18 and Rrm1
RT proteins for degradation and stops unwanted diploidization in fission
RT yeast."
RL Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell J., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Hunkle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicikert G., Aert R., Robben J., Gymnopoulos B.,
RA Weljens I., Vansteede E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga K.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.C., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."

```

```

RL Nature 415:871-880(2002).
RN [4]
RP FUNCTION, AND SUBUNIT.
RC STRAIN=972;
RX MEDLINE=99144318; PubMed=9990507;
RA Komatsu K.-I., Ochotorena I., Toda T.;
RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
RT complexes together with cullin-1 in fission yeast SCF (Skip-cullin-1-
RT F-box) ubiquitin ligase."
RL Genes Cells 3:721-735(1998).
CC -1- FUNCTION: Involved in maintenance of ploidy through proteasome
CC dependent degradation of CDK inhibitor rum1 and S-phase initiator
CC cdc18. Functions as a recognition factor for rum1 and cdc18, which
CC are subsequently ubiquitinated and targeted to the 26S proteasome
CC for degradation.
CC -1- SUBUNIT: Homodimer and heterodimer with pop1. Binds to pcu1 and
CC cdc18.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF038867; AAB95480.1; -.
DR EMBL: AF064515; AAC39496.1; -.
DR EMBL: Z98602; CAB11275.1; -.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 3.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 6.
DR PROSITE: PS00181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT DOMAIN 236 283 F-BOX.
FT REPEAT 389 417 WD 1.
FT REPEAT 429 473 WD 2.
FT REPEAT 505 533 WD 3.
FT REPEAT 545 575 WD 4.
FT REPEAT 587 615 WD 5.
FT REPEAT 625 654 WD 6.
SQ SEQUENCE 703 AA; 79605 MW; 0336A0568C152396 CRC64;

Query Match
Best local Similarity 48.1%; Score 75; DB 1; Length 703;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 2 GHKRGIAQYRDRLVVGSSDNTIRLW 29
DQ 389 GHEGVAWKXHTGNTLVGTSTDRVRW 416

RESULT 13
SE10_CAEEL STANDARD; PRT; 579 AA.
ID SE10_CAEEL
AC 093794;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sel-10 protein.
GN SEL-10 OR F55B12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```

OC Rhabditiidae: Peloderinae: Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Sims M.;  
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP REVISIONS.  
 RA Jones S.J.M.;  
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD\_REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: Z79757; CAB02129.1;  
 DR WormPep: F55B12.3; CE16120.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 5.  
 DR SMART: SM00320; WD40; 7.  
 DR SMART: SM00256; FBOX; 1.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS; 1; 5.  
 DR PROSITE: PS50082; WD\_REPEATS; 2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT DOMAIN 113 159 F-BOX.  
 FT REPEAT 245 274 WD 1.  
 FT REPEAT 286 316 WD 2.  
 FT REPEAT 328 356 WD 3.  
 FT REPEAT 368 396 WD 4.  
 FT REPEAT 408 438 WD 5.  
 FT REPEAT 453 481 WD 6.  
 FT REPEAT 493 522 WD 7.  
 FT REPEAT 579 642 WD 7.  
 FT REPEAT 642 725 WD 7.  
 SQ SEQUENCE 579 AA; 64275 MW; 0647245AF5964663 CRC64;  
 Query Match 47.8%; Score 74.5; DB 1; Length 579;  
 Best Local Similarity 44.8%; Pred. No. 0.00059;  
 Matches 13; Conservative 8; Mismatches 7; Indels 1; Gaps 1;  
 OY 2 GHK-RGIACIQYRDRLVSGSSDNTIRLM 29  
 DB 245 GHEDHVTICMQIHDDVLTGSDNTLKWV 273  
 RESULT 14  
 YK16\_YEAST  
 ID YK16\_YEAST STANDARD; PRT; 659 AA.  
 AC P36130;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 74.7 kDa Trp-Asp repeats containing protein in DAL80-GAP1  
 DE intergenic region.  
 GN YK0306.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Uristarazu L.A., Jauniaux J.-C.;

RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 6 WD\_REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: Z28261; CA882110.1;  
 DR PIR: S38108; S38108.  
 DR TRANSFAC: T04326;  
 DR SGD: S0001744; YK036C.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 1.  
 DR SMART: SM00320; WD40; 5.  
 DR PROSITE: PS00678; WD\_REPEATS; 1; 4.  
 DR PROSITE: PS50082; WD\_REPEATS; 2; 5.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 333 364 WD 1.  
 FT REPEAT 376 407 WD 2.  
 FT REPEAT 438 466 WD 3.  
 FT REPEAT 503 531 WD 4.  
 FT REPEAT 543 571 WD 5.  
 FT REPEAT 582 610 WD 6.  
 FT REPEAT 610 640 WD 6.  
 SQ SEQUENCE 659 AA; 74709 MW; 13959354C56DD37 CRC64;  
 Query Match 46.8%; Score 73; DB 1; Length 659;  
 Best Local Similarity 42.9%; Pred. No. 0.0012;  
 Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;  
 OY 2 GHKRGIAQLYRDRLVSGSSDNTIRLM 29  
 DB 543 GHDTGITSILKFKSEKLVGSDMNSVRIW 570  
 RESULT 15  
 PRP5\_SCHPO  
 ID PRP5\_SCHPO STANDARD; PRT; 473 AA.  
 AC O13615;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE pre-mRNA splicing protein prp5.  
 GN PRP5 OR CWFI OR SPBP22H7.07 OR P1024.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RA MEDLINE=99340235; PubMed=10409726;  
 RA McDonald W.H., Ohl R., Smelkova N., Frendewey D., Gould K.L.;  
 RT "Myb-related fission yeast cdc5p is a component of a 40S  
 RT snRNP-containing complex and is essential for pre-mRNA splicing";  
 RL MOL. CELL. BIOL. 19:5352-5362(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=20089027; PubMed=10620777;  
 RA Machida M., Yamazaki S., Kunihito S., Tanaka T., Kishida N., Jinno K.,  
 RA Haikawa Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,  
 RA Sakai M., Aoki K., Ogura K., Kudo Y., Kikuchi H., Zhang M.Q.,  
 RA Yanagida M.;  
 RT "A 38 kb segment containing the cdc2 gene from the left arm of fission  
 RT yeast chromosome II: sequence analysis and characterization of the  
 RT genomic DNA and cDNAs encoded on the segment.";

Yeast 16:71-80(2000).

[3]  
SEQUENCE FROM N.A.

RC STRAIN=972;

RA MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgueros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,

RA Holroyd S., Hornsby T., Howarth N., Harris D., Hidalgo J., Hodgson G.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,

RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Filtz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer T., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Spakovski G.V., Ussey D., Barrell B.G., Nurse P.,

RA "The genome sequence of Schizosaccharomyces pombe."

RA Nature 415:871-880(2002).

CC - FUNCTION: Required for both cell cycle progression at G2/M

CC and pre-mRNA splicing.

CC - SUBUNIT: INTERACTS WITH CDC5.

CC - SUBCELLULAR LOCATION: Nuclear (Probable)

CC - SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC - SIMILARITY: BELONGS TO THE PRL1/PRL2 FAMILY OF WD-REPEAT PROTEINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AF251148; AAC01399.1; -

DR EMBL; AB004535; BAA21403.1; -

DR EMBL; AL590883; CAC37375.1; -

DR InterPro: IPR001680; WD40.

DR Pfam: PF004400; WD40; 7.

DR PRINTS; PR00320; GPROTEINBRPT.

DR PRODom; PD000018; WD40; 4.

DR SMART; SM00320; WD40; 7.

DR PROSITE; PS00678; WD\_REPEATS\_1; 2.

DR PROSITE; PS00082; WD\_REPEATS\_2; 5.

DR PROSITE; PS0294; WD\_REPEAT\_REGION; 1.

DR Cell cycle; Repeat; WD repeat; mRNA processing; mRNA splicing;

DR Nuclear protein.

KW

REPEAT 161 191 WD 1.

REPEAT 203 233 WD 2.

REPEAT 245 275 WD 3.

REPEAT 287 317 WD 4.

REPEAT 329 358 WD 5.

REPEAT 370 399 WD 6.

REPEAT 419 449 WD 7.

SEQUENCE 473 AA; 52418 MW; 941F3B31A8C19799 CRC64;

Query Match 46.2%; Score 72; DB 1; Length 473;

Best Local Similarity 54.8%; Pred. No. 0.0012;

Matches 17; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

Db 286 SGHKTVA5LAVOEPDPOVWTSMDSTIRLM 316

Search completed: April 10, 2003, 13:19:37  
Job time : 3.08404 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 ; Search time 6.85262 Seconds  
(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_427\_455  
Perfect score: 156  
Sequence: 1 NGHRRGACLOXRRLVVGSSDNTIRLM 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	156	100.0	252	11	0922C7 mus musculus
2	156	100.0	563	11	0923H0 mus musculus
3	156	100.0	569	11	0921G7 mus musculus
4	156	100.0	569	11	092159 mus musculus
5	156	100.0	569	11	092159 mus musculus
6	152	97.4	510	5	044382 drosophila
7	152	97.4	510	5	09VDE3 drosophila
8	143	91.7	430	5	09B554 drosophila
9	136	87.2	122	13	08U0N3 xenopus lae
10	83	53.2	1326	5	09VZP4 drosophila
11	81	51.9	280	5	08STZ5 encephalito
12	81	51.9	730	5	08STO2 drosophila
13	80	51.3	196	13	08G6U0 melanochrom
14	80	51.3	196	13	08G6T9 asatocollap
15	80	51.3	196	13	08G6T8 labidochrom
16	80	51.3	196	13	08G6T7 tropheus du

17	80	51.3	196	13	08G6T6	08G6T6 spathodus e
18	80	51.3	196	13	08G6T5	08G6T5 cyprichromi
19	80	51.3	196	13	08G6T4	08G6T4 neolamprolo
20	80	51.3	196	13	08G6T2	08G6T2 allolampiro
21	80	51.3	196	13	08G6T1	08G6T1 neolampiro
22	80	51.3	196	13	08G6T0	08G6T0 steatocranu
23	80	51.3	196	13	08G6S9	08G6S9 tilapia but
24	80	51.3	196	13	08G6S8	08G6S8 oreochromis
25	80	51.3	423	5	08S5J1	08S5J1 encephalito
26	80	51.3	808	4	09G6T8	09G6T8 homo sapien
27	79	50.6	196	13	08G6T4	08G6T4 xenolapia
28	79	50.6	415	10	048679	048679 arabidopsis
29	79	50.6	438	10	09M3B4	09M3B4 arabidopsis
30	79	50.6	553	4	09N0X6	09N0X6 homo sapien
31	79	50.6	561	4	09G6R12	09G6R12 homo sapien
32	79	50.6	589	4	09G6LE0	09G6LE0 homo sapien
33	79	50.6	627	4	09G6A16	09G6A16 homo sapien
34	79	50.6	629	11	08VHP4	08VHP4 mus musculu
35	79	50.6	629	11	08VBP4	08VBP4 mus musculu
36	79	50.6	707	4	09G69H0	09G69H0 homo sapien
37	78.5	50.3	558	3	043071	043071 schizosach
38	77	49.4	38	13	08U0N4	08U0N4 xenopus lae
39	76	48.7	465	10	048716	048716 arabidopsis
40	76	48.7	1609	5	09G6RC3	09G6RC3 caenorhabdi
41	76	48.7	1893	5	017887	017887 caenorhabdi
42	74.5	47.8	585	5	0952T0	0952T0 caenorhabdi
43	74.5	47.8	587	5	044083	044083 caenorhabdi
44	74	47.4	47	13	08U0N5	08U0N5 xenopus lae
45	73	46.8	296	5	09VXJ1	09VXJ1 drosophila

## ALIGNMENTS

## Result 1

ID 0922C7 PRELIMINARY: PRT: 252 AA.

AC 0922C7: 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Similar to f-box and WD-40 domain protein 1B (Fragment).  
GN FBXN1B.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
DR EMBL: BC008552; AA008552.1; -;  
DR MGI: 2144023; FbxN1B.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 6.  
DR PRINTS: PR003320; GPROTEINBRT.  
DR PRODOM: PD000018; WD40; 3.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN-5.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
FT NON\_TER 1  
SQ SEQUENCE 252 AA; 28424 MW; F71737C8D7A9F75F CRC64;  
Query Match 100.0%; Score 156; DB 11; Length 252;  
Best Local Similarity 100.0%; Pred. No. 4e-16;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHRRGACLOXRRLVVGSSDNTIRLM 29  
DB 110 NGHRRGACLOXRRLVVGSSDNTIRLM 138

RESULT 2	0923H0	PRELIMINARY;	PRT;	563 AA.
ID	0923H0			
AC	0923H0;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	F-box/WD40 repeat-containing protein HOS.			
GN	FBXW1B.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N;			
RA	Bhatia N., Hetter J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;			
RT	"Mouse homolog of HOS (mHOS) is overexpressed in skin tumors and implicated in constitutive activation of NF-kappaB."			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
DR	-1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).			
DR	EMBL: AY080879; AAK72095.1; -			
DR	MCD; M61:2144023; FBxw1b.			
DR	InterPro: IPR001810; F-box.			
DR	InterPro: IPR001680; WD40.			
DR	Pfam: PF00646; F-box; 1.			
DR	Pfam: PF00400; WD40; 7.			
DR	ProDom: PD000018; WD40; 4.			
DR	PROSITE: PS00181; FBOX; 1.			
DR	PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.			
DR	PROSITE: PS50082; WD_REPEATS_2; 7.			
DR	PROSITE: PS50294; WD_REPEATS_REGION; 1.			
KW	Repeat; WD repeat.			
SQ	SEQUENCE 563 AA; 64741 MW; 9AB562F3FF5E3496 CRC64;			
Query Match	100.0%; Score 156; DB 11;			
Best Local Similarity	100.0%; Pred. No. 9, 6e-16;			
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Db	1 NCRRGIACLOQYRDRLVSGSSDNTIRLW 29			
	421 NCRRGIACLOQYRDRLVSGSSDNTIRLW 449			
RESULT 3				
09R1G7				
ID	09R1G7	PRELIMINARY;	PRT;	569 AA.
AC	09R1G7;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Beta-transducin repeat-containing protein.			
GN	BTRC.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Winston J., Elledge S.J., Harper J.W.;			
RT	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).			
DR	EMBL: A1110396; AADA1025.1; -			
DR	MCD; MGI:1338871; Btrc.			
DR	InterPro: IPR001810; F-box.			
DR	InterPro: IPR001680; WD40.			
DR	Pfam: PF00646; F-box; 1.			
DR	Pfam: PF00400; WD40; 7.			
DR	PROSITE: PS00320; GPROTEINRPT.			
DR	ProDom: PD000018; WD40; 4.			
DR	SMART; SM00326; FBOX; 1.			
DR	SMART; SM00320; WD40; 7.			

```

DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D51D9D CRC64;

Query Match
Best Local Similarity 100.0%; Score 156; DB 11; Length 569;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKRGIAQLQYRDRLVVGSSSDNTIRLM 29
|||||
Db 427 NGHKRGIAQLQYRDRLVVGSSSDNTIRLM 455

RESULT 4
O92159 PRELIMINARY; PRT; 569 AA.
ID O92159
AC O92159;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Beta-transducin repeat containing protein.
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99145465; PubMed=9990853;
RA Spencer E., Jiang J., Chen Z.J.;
RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT Slmb/beta-trcp.";
RL Genes Dev. 13:284-294(1999).
CC 1. SIMILARITY: CONTAINS 7 WD_REPEATS (TRP-ASP DOMAINS).
DR EMBL; AF112979; AAD04181.1; -.
DR MGd; MG1:1338871; BTRC.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 4.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;

Query Match
Best Local Similarity 100.0%; Score 156; DB 11; Length 569;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKRGIAQLQYRDRLVVGSSSDNTIRLM 29
|||||
Db 427 NGHKRGIAQLQYRDRLVVGSSSDNTIRLM 455

RESULT 5
O90015 PRELIMINARY; PRT; 569 AA.
ID O90015
AC O90015;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Ubiquitin ligase FWD1 (beta-transducin repeat containing protein)
DE (F-box-WD40 repeat protein 1).
GN BTRC OR FBXW1
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99145465; PubMed=9990853;
RA Spencer E., Jiang J., Chen Z.J.;
RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT Slmb/beta-trcp.";
RL Genes Dev. 13:284-294(1999).
CC 1. SIMILARITY: CONTAINS 7 WD_REPEATS (TRP-ASP DOMAINS).
DR EMBL; AF112979; AAD04181.1; -.
DR MGd; MG1:1338871; BTRC.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 4.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;

Query Match
Best Local Similarity 100.0%; Score 156; DB 11; Length 569;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKRGIAQLQYRDRLVVGSSSDNTIRLM 29
|||||
Db 427 NGHKRGIAQLQYRDRLVVGSSSDNTIRLM 455

```



OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9119275; PubMed=10097128;  
 RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,  
 RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,  
 RA Nakayama K.-I.;  
 RT "ubiquitin-dependent degradation of IkappaBalpha is mediated by a  
 RT ubiquitin ligase Skp1/Cul1/F-box protein FMDL";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99075339; PubMed=9859996;  
 RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
 RT "Identification of the receptor component of the IkappaBalpha-  
 RT ubiquitin ligase";  
 RL Nature 396:590-594(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBD databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV;  
 RX MEDLINE=21601157; PubMed=11735228;  
 RA Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,  
 RA Nakayama K.-I.;  
 RT "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of  
 RT Caenorhabditis elegans SEL-10";  
 RL Genomics 78:214-222(2001).  
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF081887; AAD17755.1; -;  
 DR EMBL: AF099932; AAD08701.1; -;  
 DR EMBL: BC003989; AAH03989.1; -;  
 DR EMBL: AF391190; AAL40929.1; JOINED.  
 DR EMBL: AF391178; AAL40929.1; JOINED.  
 DR EMBL: AF391179; AAL40929.1; JOINED.  
 DR EMBL: AF391180; AAL40929.1; JOINED.  
 DR EMBL: AF391181; AAL40929.1; JOINED.  
 DR EMBL: AF391182; AAL40929.1; JOINED.  
 DR EMBL: AF391183; AAL40929.1; JOINED.  
 DR EMBL: AF391184; AAL40929.1; JOINED.  
 DR EMBL: AF391185; AAL40929.1; JOINED.  
 DR EMBL: AF391186; AAL40929.1; JOINED.  
 DR EMBL: AF391187; AAL40929.1; JOINED.  
 DR EMBL: AF391188; AAL40929.1; JOINED.  
 DR EMBL: AF391189; AAL40929.1; JOINED.  
 DR MGI: MGI:1338871; Btgc.  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ligase; Repeat; WD repeat.  
 SQ SEQUENCE 369 AA; 65105 MW; BC7D6544815H2296 CRC64;

Query Match 100.0%; Score 156; DB 11; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 9, 7e-16;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NGHKGIACTQYRDRLVYSGSSDNTIRLM 29  
 |||

DB 427 NGHKGIACTQYRDRLVYSGSSDNTIRLM 455  
 RESULT 6  
 ID 044382 PRELIMINARY; PRT; 510 AA.  
 AC 044382;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE SLMB.  
 GN SLMB OR SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9812115; PubMed=9461217;  
 RA Jiang J., Struhl G.;  
 RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-  
 RT box/WD40-repeat protein Slmb";  
 RL Nature 391:493-496(1998).  
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF032878; AAC38852.1; -;  
 DR FLYBASE: FBgn0023423; slmb.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 510 AA; 58952 MW; DDB0243D3730A5E8 CRC64;

Query Match 97.4%; Score 152; DB 5; Length 510;  
 Best Local Similarity 96.6%; Pred. No. 3, 6e-15;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 NGHKGIACTQYRDRLVYSGSSDNTIRLM 29  
 |||

DB 370 NGHKGIACTQYRDRLVYSGSSDNTIRLM 398

RESULT 7  
 ID 09VDE3 PRELIMINARY; PRT; 510 AA.  
 AC 09VDE3;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE SLMB protein (SLMB).  
 GN SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Asgari A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkov D., Bouchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gottlieb J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Honck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Lasko B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Massarman D.A., Welnsch G.M., Welnschbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.,  
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RU Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVARY, AND IMAGINAL DISC;  
 RX MEDLINE=20245299; PubMed=10781936;  
 RA Miletich I., Limbourg-Bouchon B.;  
 RT "Drosophila null limb clones transiently deregulate Hedgehog-  
 independent transcription of wingless in all limb discs and induce  
 decapentaplegic transcription linked to imaginal disc regeneration.";  
 RL Mech. Dev. 93:15-26(2000).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AE003733; AAF5853.1; -  
 DR EMBL; AF22924; AAF6324.1; -  
 DR EMBL; AF22923; AAF6323.1; -  
 DR Flybase; FBgn0023423; slmb.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00646; F-box; 1.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KM Repeat; WD repeat.  
 SQ SEQUENCE 510 AA; 58952 MW; FADSDFI26F58A012 CRC64;

Query Match 97.4%; Score 152; DB 5; Length 510;  
 Best local Similarity 96.6%; Pred. No. 3.6e-15;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGHKGIACTQYRDLRVVSGSSDNTIRLM 29  
 DB 370 NGHKGIACTQYRDLRVVSGSSDNTIRLM 398

RESULT 8  
 Q9BJ54

ID Q9BJ54 PRELIMINARY; PRT; 430 AA.  
 AC Q9BJ54;  
 DT 01-JUN-2001 (TRENBLREL. 17, Created)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
 DE 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Beta-transducin repeat-containing protein (Fragment).  
 OS Heterodera glycyines (Soybean cyst nematode).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;  
 OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.  
 OX NBI\_TaxID=51029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kovalova E.S., Yakovlev A.G., Master E.P.,  
 RT "Plant parasitic nematode b-trcp.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AF339101; AAK26376.1; -  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KM Repeat; WD repeat.  
 FT NON\_TER  
 SQ SEQUENCE 430 AA; 47916 MW; 4ABC3FF2DFE3A50B CRC64;

Query Match 91.7%; Score 143; DB 5; Length 430;  
 Best local Similarity 89.7%; Pred. No. 7.6e-14;  
 Matches 26; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGHKGIACTQYRDLRVVSGSSDNTIRLM 29  
 DB 251 NGHKGIACTQYRDLRVVSGSSDNTIRLM 279

RESULT 9  
 ID Q9BJ54 PRELIMINARY; PRT; 122 AA.  
 AC Q9BJ54;  
 DT 01-MAR-2002 (TRENBLREL. 20, Created)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
 DE 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Beta-Trcp protein (Fragment).  
 GN BETA-Trcp.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Carnevali F.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ballarino M.;  
 RT "Analisi strutturale e funzionale del gene beta-trcp in *Xenopus*  
 laevis";  
 RL Thesis (2001).  
 RL Department of Genetics and Molecular Biology "Charles Darwin",  
 RL University of Rome La Sapienza, Rome, Italy.  
 DR EMBL; AJ428938; CAD21935.1; -  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 3.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR ProDom; PD000018; WD40; 2.  
 DR SMART; SM00320; WD40; 3.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_2.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 3.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.

FT NON\_TER 1 1  
 FT NON\_TER 122 122  
 SQ SEQUENCE 122 AA; 13682 MW; 36BB6B39AC8F5387 CRC64;

Query Match 87.2%; Score 136; DB 13; Length 122;  
 Best Local Similarity 100.0%; Pred. NO. 2.4e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGHRRGIACQYRDRLVSSGSSDNTI 26  
 Db 97 NGHRRGIACQYRDRLVSSGSSDNTI 122

RESULT 10  
 O9VZF4 PRELIMINARY; PRT; 1326 AA.  
 AC O9VZF4:  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE CG15010 protein (LD21322p) (LD30271p).  
 OS AGO OR CG15010.  
 GN Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Ephydrota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Butkus K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Talali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris D., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celiker S.;  
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nuno J., Pacle J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celiker S.;  
 RL Submitted (Jan-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- STIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAIN).  
 DR EMBL; AF003480; AAG22246.1; -;  
 DR EMBL; AY061300; AAL28848.1; -;  
 DR EMBL; AY075401; AAL68231.1; -;  
 DR FlyBase; FBgn0041171; ago.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00646; F-box; 1.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR PRODOM; PD000018; WD40; 2.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS0181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 1326 AA; 141360 MW; 3F42C873CFA3027F CRC64;

OY 1 NGHRRGIACQYRDRLVSSGSSDNTIRLM 29  
 Db 1235 NKGHRAVTCIQFNSRFVYTSSDDCTVKLM 1263

Query Match 53.2%; Score 83; DB 5; Length 1326;  
 Best Local Similarity 44.8%; Pred. NO. 0.00059;  
 Matches 13; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

RESULT 11  
 O8STZ5 PRELIMINARY; PRT; 280 AA.  
 AC O8STZ5:  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Hypothetical protein ECUI1\_1970.  
 GN ECUI1\_1970.  
 OS Encephalitozoon cuniculi.  
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.  
 OX NCBI\_TaxID=6035;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BB-M1;  
 RA Genoscope;  
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BB-M1;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
 RA Prestier G., Barbe V., Peyretallade E., Brotlier P., Wincker P.,  
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissbach J., Vivares C.P.;  
 RT "Genome sequence and gene compaction of the eukaryote parasite  
 Encephalitozoon cuniculi."  
 RL Nature 414:450-453(2001).  
 DR EMBL; AL590450; CAD26107.1; -;

KW Hypothetical protein.  
SQ SEQUENCE 280 AA; 31339 MW; 501767DF5970F538 CRC64;

Query Match	51.9%;	Score 81;	DB 5;	Length 280;
Best Local Similarity	53.6%;	Pred. No. 0.00022;		
Matches 15;	Conservative 3;	Mismatches 10;	Indels	

QY 2 GHKRGIAQLQYRDLVAGSSSDNTIRLW 29  
|| : || | : ||||| ||| : |  
Db 94 GHMSVVLCLDYDRIRLFSSSDCTIRW 121

RESULT 12  
Q8ST02

DT 01-JUN-2002 (TREMBlEt. 21, last sequence update)  
DT 01-JUN-2002 (TREMBlEt. 21, last annotation update)  
DE Hypothetical 54.5 kDa Trp-ASP repeats containing protein ZC302.2  
OS Dictyostelium discoideum (slime mold)  
US Eukaryota; Mycetesozoa; Dictyostellida; Dictyostellium;  
NCBI\_TaxID=44689;

RA Gloeckner G., Eichlinger L., Szatranski K., Pachebat J., Dear P.,  
 RA Lehmann R., Baumgart C., Parra G., Arill J. F., Guigo R., Kumpf K.,  
 RA Tunga B., Cox E., Quail M. A., Platner M., Rosenblatt A., Noegel A. A.,  
 RA "Sequence and Analysis of Chromosome 2 of Dictyostelium".  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC115681; AAL92670.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 730 AA; 80830 MW; 9E0D5A8B373E59E7 CRC64;

Query Match	51.98;	Score 81;	DB 5;	Length 730;
Best Local Similarity	45.28;	Pred. No. 0.00063;		
Matches	14;	Conservative	8;	Mismatches 7;
				Indels

```

0Y      1 NGHKGICAIQY--RDRLVGGSSDNTIKLM 29
      :|||: :|||: :|||: :|||: :|||: :|||:
Db      372 SGHKRNKVCVEFIINGLTLAGSSDNTIKLM 402

```

RESULT 13  
Q8QGU0  
ID Q8QGU0 PRELIMINARY; PRT: 196 AA.

01-JUN-2002 (TREMblrel\_21, Created)  
01-JUN-2002 (TREMblrel\_21, Last sequence update)  
01-JUN-2002 (TREMblrel\_21, Last annotation update)  
F-box/WD40-repeat protein (Fragment).  
HAGOROMO.  
Melanochromis auratus (Golden cichlid).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Labroidae;  
Cichlidae; Melanochromis.  
NCBI\_TaxId=27751;  
11

SCIENCE FROM N.A.  
MEDLINE-21918437; PubMed-11919300;  
Terao Y., Morikawa N., Kawakami K., Okada N.:  
"Accelerated Evolution of the Surface Amino Acids in the WD-Repeat  
Domain Encoded by the hachoro Gene in an Explosively Speciated  
Lineage of East African Cichlid Fishes.",  
Mol. Biol. Evol. 19:574-578(2002).  
PMID: 120075632; DOI:10.1093/molbev/19.5.574

SEQUENCE	196 AA; 21854 MW; 499A27323286108F7 CRC64;
Query Match	51.3%; Score 80; DB 13; Length 196;

Best Local Similarity 41.4%; Pred. No. 0.00021;  
Matches 12; Conservative 8; Mismatches 0; Indels 0; Gaps 0

QY 1 NHHKRGIAQLQYRDLVVGSSDNTIRLW 29  
::||:::|:|:|:|:|:|:|:  
Db 34 SGHNOEVNCCVDAKDGLIISGRSDRTARIW 62

RESULT 14  
080GT9  
ID 080GT9  
PRELIMINARY: PBT. 196 AA

DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DF 01-JUN-2002 (TrEMBLrel. 21, last sequence update)  
 DI 01-JUN-2002 (TrEMBLrel. 21, last annotation update)  
 DE F-box/Wd40-repeat protein (Fragment).  
 GN HAO6ROMO.  
 OS Asatolatolliapia brownae.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;  
 OC Cichlidae; Asatolatolliapia.  
 OX NCBI\_TaxID=64538;  
 RN [1]

RX MEDLINE=21918437; PubMed=11919300;  
 RA Terai Y., Morikawa N., Kakakami K., Okada N.;  
 RT "accelerated Evolution of the Surface Amino Acids in the WD-Repeat  
 PT Domain Encoded by the hagoromo Gene in an Explosively Speciated  
 RT Lineage of East African Cichlid Fishes.";  
 RL Mol. Biol. Evol. 19:574-578(2002).  
 DR EMBL; AB075464; BAB85689.1; -.  
 RL NON\_TER 1  
 FT NON\_TER 196  
 FT 196  
 SQ SEQUENCE 196 AA; 21870 MM; 5D272C73336108E7 CRC64;

Query Match	51.3%	Score 80;	DB 13;	Length 196;.
Best Local Similarity	41.4%	Pred. No. 0.00021;		
Matches 12; Conservative	8;	Mismatches 9;	Indels 0;	Gaps 0;

QY    1 NGHKRGIACLÖYRDRLVSGSNDTIRLM 29  
      :| | : : | : |||| | | : |  
Db    34 SGHNÖEVNCVDKDKGLITSGSRDRTARIW 62

RESULT 15  
Q08QGT8  
ID Q08QGT8 PRELIMINARY; PRT; 196 AA.

DT 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)  
DE F-box/WD40-repeat protein (Fragment).  
GN HAGROMO.  
OS  
OS Laidichromis caeruleus (blue streak hap)  
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
OC Actinopterygii: Neopterygii: Teleostei: Euteleostei: Neoteleostei:  
OC Acanthomorphia: Acanthopterygii: Percormorpha: Perciformes: Labroidae:  
OC Cichlidae: Laidichromis.  
OX NCBI\_TaxId=50897;  
ID 111

RT MEDLINE=21918437; PubMed=11919300;  
RT Teraï Y., Morikawa N., Kawakami K., Okada N.;  
RT "Accelerated Evolution of the Surface Amino Acids in the WD-Repeat  
RT Domain Encoded by the hagoromo Gene in an Explosively Speciated  
RT Lineage of East African Cichlid Fishes."; *Mol. Biol. Evol.* 19:574-578(2002).  
PMID: 12007545; PubMed=11905742;  
RT FMR1. ABO75455.1; PubMed=11905742;

FT	NON_LEN	1	1
NON_TER	196	196	
SEQUENCE	196 AA; 21914 MW; BC2A68399CA1246 CRC64;		

Fri, Apr 11 13:31:12 2003

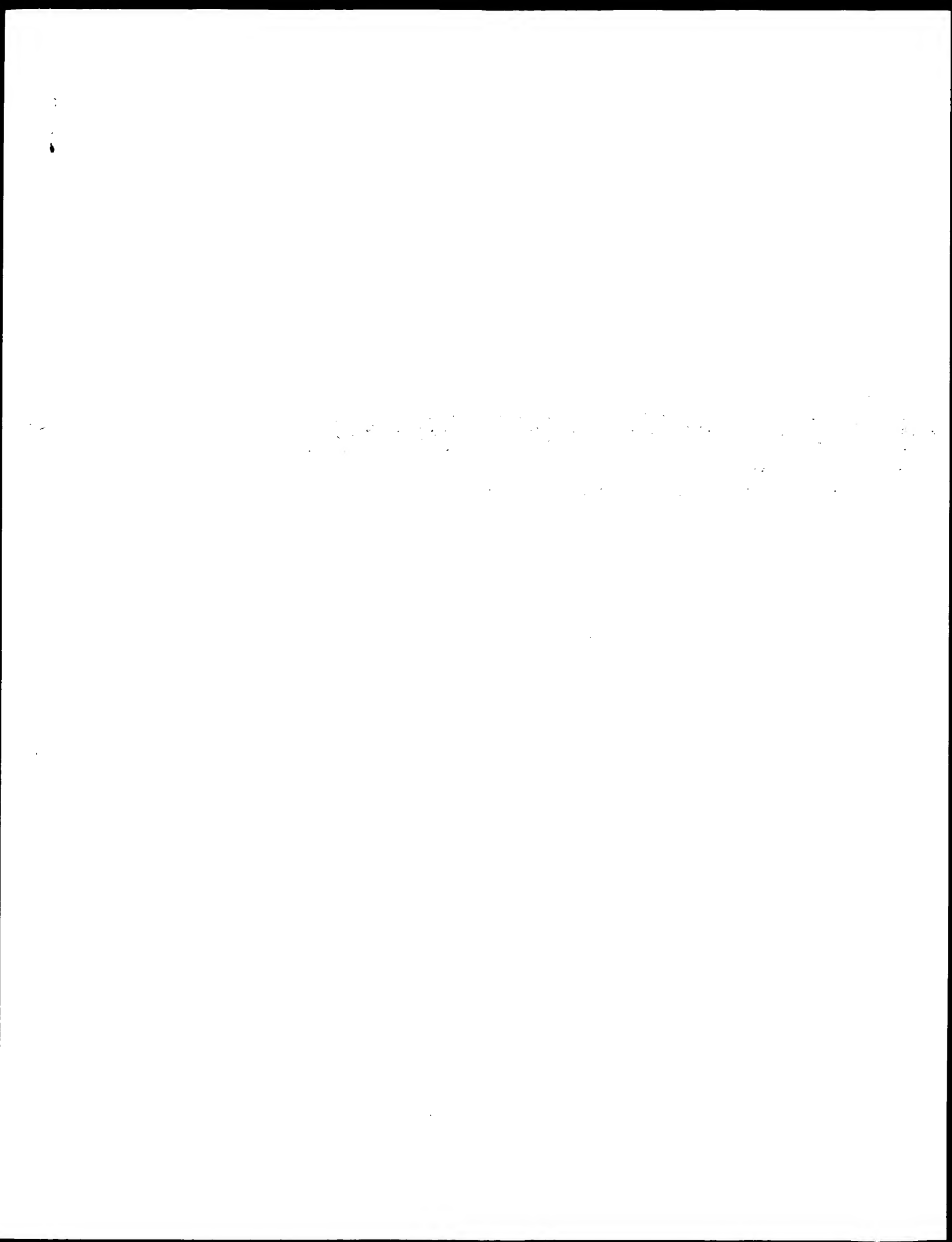
us-09-601-168b-2\_copy\_427\_455.rspt

Page 7

Query Match 51.3%; Score 80; DB 13; Length 196;  
Best Local Similarity 41.4%; Pred. No. 0.00021;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

```
QY      1 NGHKRGIAcLOYRDRLVSSGSSDNTIRLW 29
      :||: :|: :||: ||| | | | | |
Db      34 SGHNQEVNCVDPAKDGLLISGRDRTARIW 62
```

Search completed: April 11, 2003, 11:51:52  
Job time : 14.8526 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:20:44 : Search time 8.01218 Seconds  
(without alignments)  
432.406 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_467\_492  
Perfect score: 138  
Sequence: 1 EGHETLVRCIFRDNMRIYSGAYDGKI 26

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :  
1: A-Geneseq.101002.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	100.0	517	AA85852	WD-40 domain-contg
2	138	100.0	542	AA96696	Human E3 ubiquitin
3	138	100.0	542	AA79127	Human protein SEO
4	138	100.0	542	AA40208	Human polypeptide
5	138	100.0	550	AA41994	Human polypeptide
6	138	100.0	569	AA24054	A human beta-trans
7	138	100.0	569	AA12812	Mouse ubiquitin 11
8	138	100.0	569	AA12813	Human beta-transdu
9	138	100.0	569	AA96697	Human beta-TFPC
10	138	100.0	569	AA783041	F-box protein FBP-

11	138	100.0	569	AA83250	F-box protein hbet
12	138	100.0	569	AA83254	F-box protein FWD1
13	138	100.0	569	AA44249	Human cell signal1
14	138	100.0	569	AA88238	Human F211 protein
15	138	100.0	579	AA78583	Human protein SEO
16	138	100.0	590	AA00847	Human bone marrow
17	138	100.0	605	AA78582	Human protein SEO
18	138	100.0	608	AA00960	Human bone marrow
19	138	100.0	632	AA78584	Human protein SEO
20	138	100.0	654	AA79566	Human protein SEO
21	138	100.0	654	AA79567	Human protein SEO
22	138	100.0	654	AA79568	Human protein SEO
23	133	96.4	29	AA84940	Peptide RY from a
24	132	95.7	510	AB59857	Drosophila melanog
25	82	59.4	540	AA72465	Human hippocampal
26	82	59.4	540	AA72466	Human mammary sel-
27	82	59.4	540	AA72467	Human mammary sel-
28	82	59.4	545	AA72468	Human mammary sel-
29	82	59.4	545	AA72469	Human mammary sel-
30	82	59.4	545	AA72470	Human mammary sel-
31	82	59.4	553	AA72471	Human mammary sel-
32	82	59.4	553	AA72472	Human mammary sel-
33	82	59.4	553	AA72473	Human mammary sel-
34	82	59.4	559	AA72474	Human mammary sel-
35	82	59.4	559	AA72475	Human mammary sel-
36	82	59.4	589	AA72476	Human mammary sel-
37	82	59.4	589	AA72477	Human mammary sel-
38	82	59.4	589	AA72478	Human mammary sel-
39	82	59.4	589	AA72479	Human mammary sel-
40	82	59.4	592	AA72480	Human mammary sel-
41	82	59.4	592	AA72481	Human mammary sel-
42	82	59.4	626	AA72482	Human mammary sel-
43	82	59.4	627	AA72483	Human mammary sel-
44	82	59.4	627	AA72484	Human mammary sel-
45	82	59.4	666	AA72485	Human C-term mychi

## ALIGNMENTS

RESULT 1  
AA85852 standard; peptide: 517 AA.  
ID AA85852 standard; peptide: 517 AA.  
XX  
AC AA85852; (first entry)  
XX  
DT 13-SEP-1996  
XX  
DE WD-40 domain-contg. beta-TFPC protein.  
XX  
KW WD40 repeat region; beta-transducin; protein-protein interaction; drug;  
KW Intracellular signalling; protein kinase C; homology; motif; modulator;  
KW receptors of activated protein kinase; enzyme; activity; isozyme; human.  
XX  
OS Synthetic.  
XX  
PN WO9521252-A2.  
XX  
PD 10-AUG-1995.  
XX  
PF 31-JAN-1995; 95WO-US01210.  
XX  
PR 01-FEB-1994; 94US-0190802.  
XX  
(STRD ) UNIV LELAND STANFORD JUNIOR.  
XX  
PI Mochly-Rosen D. Ron D;  
XX  
DR WPI; 1995-283772/37.  
XX  
PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the  
PT activity of a protein, eg. protein kinase C, which interacts with a  
PT protein contg. a WD-40 region.

XX Example 5: Page 80-82; 351pp: English.  
 PS  
 CC Proteins AAR8581-92 are protein which contain at least one WD-40 (also  
 CC called beta-transducing homologous) amino acid repeat motifs. The WD-40  
 CC regions are involved in protein-protein interactions between proteins  
 CC involved in intracellular signaling. An example of such an interaction  
 CC is between protein kinase C and receptors of activated protein kinase  
 CC (RACK), esp. RACK-1 (AAR8585). Proteins AAR8581-82 were isolated based  
 CC on homology with beta-transducin, whereas proteins AAR84928-R85063 and  
 CC isolated based on homology with the WD-40 consensus sequence (AAR85893).  
 CC The proteins were used to construct the peptides AAR84928-R85063 and  
 CC AAR85786-R85842. The peptides can be used to identify target proteins  
 CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of  
 CC proteins involved in protein-protein interaction and to screen for drugs  
 CC that will affect protein-protein interaction involving WD-40 domains.  
 SO Sequence 517 AA:  
 Query Match 100.0%; Score 138; DB 16; Length 517;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EGHELVRCIRFDPNKRIVSGAYDGKI 26  
 Db 431 EGHELVRCIRFDPNKRIVSGAYDGKI 456  
 RESULT 2  
 ID AAY96696 standard; Protein: 542 AA.  
 XX AAY96696;  
 AC  
 XX 26-SEP-2000 (first entry)  
 DT  
 XX Human E3 ubiquitin ligase.  
 DE  
 XX E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KW anti-inflammatory; immunosuppressive; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200034447-A2.  
 PD 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99WO-US29371.  
 XX  
 PR 10-DEC-1998; 98US-0210060.  
 XX  
 PA (SIGN-) SIGNAL PHARM INC.  
 PA (YISS) YISSON RES & DEV CO.  
 XX  
 PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;  
 PI Lavon I, Yaron A;  
 PI WPI: 2000-431294/37.  
 DR N-PSDB; AAK51228.  
 XX  
 PT Polypeptide enhancing phosphorylated I-kappaB ubiquitination useful for  
 PT treating disorder associated with NF-kappaB activation e.g. cancer,  
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
 PT variant  
 XX  
 PS Claim 1; Page 70-72; 77pp: English.  
 XX  
 CC This is human E3 ubiquitin ligase (E3), which is homologous to human  
 CC beta-TrCP, an F-box/WD protein family member. E3 enhances ubiquitination  
 CC of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor  
 CC kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the  
 CC ubiquitin pathway is useful for identifying modulators of this process

CC for use in treating diseases associated with activation of NF-kappa-B. In  
 CC vitro analysis suggests that deletion of the F-box results in a protein  
 CC that functions as a dominant negative molecule in vivo. Transient  
 CC over-expression of delta-beta-TrCP (a deletion mutant) inhibited the  
 CC degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells,  
 CC resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be  
 CC used to screen for modulators of NF-kappa-B activity. E3 and beta-TrCP  
 CC can be used to modulate NF-kappa-B to treat inflammatory diseases,  
 CC autoimmune diseases, cancer and viral infections.  
 SO Sequence 542 AA:  
 Query Match 100.0%; Score 138; DB 21; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EGHELVRCIRFDPNKRIVSGAYDGKI 26  
 Db 440 EGHELVRCIRFDPNKRIVSGAYDGKI 465  
 RESULT 3  
 ID AAM79127 standard; Protein: 542 AA.  
 XX AAM79127;  
 AC  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX Human protein SEQ ID NO 1789.  
 DE  
 XX  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157190-A2.  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-FEB-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 PI WPI: 2001-476283/51.  
 DR N-PSDB; AAK52260.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20; Page 4147-4148; 6221pp: English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,



CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM60020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX Sequence 542 AA;

Query Match 100.0%; Score 138; DB 22; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGHEELVRCIRFDPNKRIVSGAYDGKI 26  
 |||||  
 Db 440 EGHEELVRCIRFDPNKRIVSGAYDGKI 465

RESULT 4  
 AAM40208  
 ID AAM40208 standard; Protein: 542 AA.

XX AAM40208;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3353.

DE Human; nootropic; immunosuppressant; cytoskeletal; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia.

XX Homo sapiens.

OS WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSEQ-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB: AAI59364.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Example 5; SEQ ID NO 3353; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, Leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 542 AA;

Query Match 100.0%; Score 138; DB 22; Length 542;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGHEELVRCIRFDPNKRIVSGAYDGKI 26  
 |||||  
 Db 440 EGHEELVRCIRFDPNKRIVSGAYDGKI 465

RESULT 5  
 AAM41994  
 ID AAM41994 standard; Protein: 550 AA.

XX AAM41994;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6925.

DE Human; nootropic; immunosuppressant; cytoskeletal; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia.

XX Homo sapiens.

OS WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSEQ-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB: AAI61150.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6925; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,



DR N-PSDB: AAA73131.  
 XX F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of Ikapab or beta-catenin  
 XX  
 XX  
 PS Claim 2; Page 9-10; 19pp; Japanese.  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of Ikapab or  
 CC beta-catenin and is constituted by Skp1 protein, Cull protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 XX Sequence 569 AA:  
 SQ  
 Query Match 100.0%; Score 138; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 EGHEELVRCIRPDNKRIVSGAYDGKI 26  
 ID AAB12813 standard; protein; 569 AA.  
 XX AAB12813;  
 XX 27-NOV-2000 (first entry)  
 DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.  
 XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikapab;  
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KW beta-TrCP.  
 XX Homo sapiens.  
 OS  
 XX JP2000166542-A.  
 PN 20-JUN-2000.  
 PD 02-DEC-1998; 98JP-0343437.  
 XX 02-DEC-1998; 98JP-0343437.  
 PR (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 PA  
 DR WPI: 2000-485550/43.  
 DR N-PSDB: AAA73132.  
 PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of Ikapab or beta-catenin  
 XX  
 XX  
 PS Claim 3; Page 10-12; 19pp; Japanese.  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of Ikapab or  
 CC beta-catenin and is constituted by Skp1 protein, Cull protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 XX

SQ Sequence 569 AA:  
 Query Match 100.0%; Score 138; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 EGHEELVRCIRPDNKRIVSGAYDGKI 26  
 ID AAB12813 standard; protein; 569 AA.  
 XX AAB12813;  
 XX 27-NOV-2000 (first entry)  
 DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.  
 XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikapab;  
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KW beta-TrCP.  
 XX Homo sapiens.  
 OS  
 XX JP2000166542-A.  
 PN 20-JUN-2000.  
 PD 02-DEC-1998; 98JP-0343437.  
 XX 02-DEC-1998; 98JP-0343437.  
 PR (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 PA  
 DR WPI: 2000-485550/43.  
 DR N-PSDB: AAA73132.  
 PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of Ikapab or beta-catenin  
 XX  
 XX  
 PS Claim 3; Page 10-12; 19pp; Japanese.  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of Ikapab or  
 CC beta-catenin and is constituted by Skp1 protein, Cull protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 XX

RESULT 9  
 AAY96697  
 ID AAY96697 standard; Protein; 569 AA.  
 XX AAY96697;  
 AC  
 XX 26-SEP-2000 (first entry)  
 DE Human beta-TrCP.  
 DE E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KW anti-inflammatory; immunosuppressive; cytostatic.  
 XX  
 OS  
 XX Homo sapiens.  
 PN WO200034447-A2.  
 PD 15-JUN-2000.  
 XX 10-DEC-1999; 99WO-US29371.  
 PF 10-DEC-1998; 98US-0210060.  
 PR (SIGN-) SIGNAL PHARM INC.  
 PA (YISS) YISSUM RES & DEV CO.  
 XX  
 PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;  
 PI Lavon I, Yaron A;  
 PI  
 DR WPI: 2000-431294/37.  
 DR N-PSDB: AAA51229.  
 PT polypeptide enhancing phosphorylated Ikapab ubiquitination useful for  
 PT treating disorder associated with NF-kappa-B activation e.g. cancer;  
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
 PT variant  
 XX  
 XX  
 PS Claim 21; Page 72-74; 77pp; English.  
 CC Human beta-TrCP, an F-box/WD protein family member, has been shown to  
 CC have homology to human E3 ubiquitin ligase (E3). E3 enhances  
 CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of  
 CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B  
 CC degradation via this process for use in treating diseases associated with  
 CC modulators of this process for use in treating diseases associated with  
 CC activation of NF-kappa-B. In vitro analysis suggests that deletion of  
 CC the F-box results in a protein that functions as a dominant negative  
 CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a  
 CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha  
 CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated  
 CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B  
 CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat  
 CC inflammatory diseases, autoimmune diseases, cancer and viral infections.  
 CC  
 XX Sequence 569 AA:  
 SQ  
 Query Match 100.0%; Score 138; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 EGHEELVRCIRPDNKRIVSGAYDGKI 26

Db 467 EGHELVRCIRFDNKRIVSGAYDCKI 492

RESULT 10

AAV83041

ID AAV83041 standard; Protein; 569 AA.

AC AAV83041;

DT 16-AUG-2000 (first entry)

DE F-box protein FBP-1.

KW F-box protein; FBP; diagnosis; treatment; screening; agonist;  
antagonist; proliferative disorder; differentiative disorder;  
breast cancer; prostate cancer; ovarian cancer; cancer;  
small cell lung carcinoma; immune disorder; cardiovascular disorder;  
inflammatory disorder; human.

OS Homo sapiens.

PN WO200012679-A1.

PD 09-MAR-2000.

PE 27-AUG-1999; 99WO-US19560.

PR 28-AUG-1998; 98US-0098355.

PR 03-FEB-1999; 99US-0118568.

PR 15-MAR-1999; 99US-0124449.

PA (UYNT ) UNITV NEW YORK STATE.

PI Chlaui DS, Pagano M, Latres E;

DR MPI: 2000-256635/22.

DR N-PSDB; AA293350.

PT Novel nucleic acid for screening compounds useful for treating

disorders comprising sequences encoding ubiquitin ligases

Disclousure; Figure 3a; 245pp; English.

CC Nucleic acids encoding substrate-targeting subunits of ubiquitin  
ligases with F-box motifs (F-box proteins) are useful for diagnosis  
of proliferative and differentiated related disorders by measuring  
CC FBP gene expression. Cells expressing such proteins or  
their fragments are useful for screening compounds. The compounds  
are agonists or antagonists, which are useful for treating a  
CC proliferative or differentiative disorder in a mammal such as  
breast, ovarian and prostate cancer and small cell lung carcinoma  
CC and also major opportunistic infections, immune disorders,  
cardiovascular diseases and inflammatory disorders. FBP protein,  
CC analogs, derivatives and their subsequences, anti-FBP antibodies  
are also useful in diagnosis of the disorders.

SO Sequence 569 AA;

Query Match 100.0%; Score 138; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 9.4e-15;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHELVRCIRFDNKRIVSGAYDCKI 26  
ID AAV83250 standard; Protein; 569 AA.

RESULT 11  
AAV83250  
ID AAV83250 standard; Protein; 569 AA.

AC AAV83250;

DT 16-AUG-2000 (first entry)

DE F-box protein hbetaTtrcp.

KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;  
destabilisation; proteolysis; drug discovery; gene therapy; cancer;  
oncoprotein; Huntington's disease; gene knockout; delivery systems;  
human.

OS Homo sapiens.

PN WO200022110-A2.

PD 20-APR-2000.

PE 08-OCT-1999; 99WO-US23705.

PR 09-OCT-1998; 98US-0103787.

PA (HARD ) HARVARD COLLEGE.

PI Zhou P, Howley P;

DR MPI: 2000-317970/27.

DR N-PSDB; AA293710.

PT Targeting degradation of polypeptide useful for treating cancer and  
other proliferative disorders, involves conjugating polypeptide with  
ubiquitin protein ligase or inhibiting ubiquitination using organic  
compound

Claim 9; Page 171; 185pp; English.

CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin  
ligases) which can be used for the targeted degradation of a target  
CC polypeptide in vivo. Targeted degradation is achieved by expressing  
CC the ubiquitin ligase in a cell linked to the interaction domain of  
CC the target polypeptide and thereby recruiting the target polypeptide  
to the ubiquitin ligase. Such methods are useful for decreasing or  
increasing the level of a target polypeptide and for creating and  
CC expressing a destabilized polypeptide which is subjected to SCF  
mediated proteolysis. Degrading any desired protein in a cell is  
CC useful for preventing or treating diseases caused by the presence of  
CC abnormal amount of the specific polypeptides, for drug discovery and  
for gene therapy. Diseases treated include cancer, by degradation of  
CC oncoproteins, Huntington's disease, other proliferative disorders and  
microbial infections. The method provides a quick and easy  
CC alternative to gene knockout technology. The target polypeptide can  
be degraded at all stages, or a specific stage, of development in the  
mature animal.

SO Sequence 569 AA;

Query Match 100.0%; Score 138; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 9.4e-15;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHELVRCIRFDNKRIVSGAYDCKI 26  
ID AAV83254 standard; Protein; 569 AA.

RESULT 12  
AAV83254  
ID AAV83254 standard; Protein; 569 AA.

DE F-box protein FWDip.

XX	Ubiquitin ligase; SCF; F-box protein; targeted degradation;	
KW	destabilisation; proteolysis; drug discovery; gene therapy; cancer;	
KM	oncoprotein; Huntington's disease; gene knockout; delivery systems;	
KW	mouse; ss.	
XX		
XX	Mus musculus.	
XX		
PN	W0200022110-A2.	
XX		
PD	20-APR-2000.	
XX		
PF	08-OCT-1999; 99W0-US23705.	
XX		
PR	09-OCT-1998; 98US-0103787.	
XX		
PA	(HARD ) HARVARD COLLEGE.	
XX		
PL	Zhou P, Howley P;	
XX		
DR	WPI: 2000-317970/27.	
XX	N-PSDB; AA93714.	
PT	Targeting degradation of polypeptide useful for treating cancer and	
PT	other proliferative disorders, involves conjugating polypeptide with	
PT	ubiquitin protein ligase or inhibiting ubiquitination using organic	
PT	compound	
XX		
PS	Claim 9; Page 184-185; 185pp; English.	
XX		
CC	The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin	
CC	ligases) which can be used for the targeted degradation of a target	
CC	polypeptide in vivo. Targeted degradation is achieved by expressing	
CC	the ubiquitin ligase in a cell linked to the interaction domain of	
CC	the target polypeptide and thereby recruiting the target polypeptide	
CC	to the ubiquitin ligase. Such methods are useful for decreasing or	
CC	increasing the level of a target polypeptide and for creating and	
CC	expressing a destabilized polypeptide which is subjected to SCF	
CC	mediated proteolysis. Degrading any desired protein in a cell is	
CC	useful for preventing or treating diseases caused by the presence of	
CC	abnormal amount of the specific polypeptides, for drug discovery and	
CC	for gene therapy. Diseases treated include cancer, by degradation of	
CC	oncoproteins, Huntington's disease, other proliferative disorders and	
CC	microbial infections. The method provides a quick and easy	
CC	alternative to gene knockout technology. The target polypeptide can	
CC	be degraded at all stages, or a specific stage, of development in the	
CC	mature animal.	
XX		
XX		
SO	Sequence 569 AA:	
	Query Match 100.0%; Score 138; DB 21; Length 569;	
	Best Local Similarity 100.0%; Pred. No. 9.4e-15;	
	Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1 EGHHELVRICIRFDNKRIVSGAYDGKI 26	
DB	467 EGHHELVRICIRFDNKRIVSGAYDGKI 492	
	RESULT 13	
	AA44249	
ID	AA44249 standard; Protein; 569 AA.	
XX		
AC	AA44249;	
XX		
DT	28-FEB-2000 (first entry)	
XX		
DE	Human cell signalling protein-12.	
XX		
KW	Cell signalling protein-12; CSIGP-12; cell proliferation;	
KW	inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;	
KW	arteriosclerosis; Addison's disease; multiple sclerosis.	

OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	Modified-site	19
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	39
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	91
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	109
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	162
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	266
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	288
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	328
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	376
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	381
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	411
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	418
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	451
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	514
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	519
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	535
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	536
FT	Modified-site	/note= "Potential phosphorylation site"
FT	Modified-site	17
FT	Modified-site	/note= "Potential glycosylation site"
FT	Modified-site	77
FT	Modified-site	/note= "Potential glycosylation site"
FT	Modified-site	416
FT	Modified-site	/note= "Potential glycosylation site"
FT	Modified-site	320..334
FT	Modified-site	/label= Signature_sequence
FT	Modified-site	360..374
FT	Modified-site	/label= Signature_sequence
FT	Modified-site	403..417
FT	Modified-site	/label= Signature_sequence
FT	Modified-site	443..457
FT	Modified-site	/label= Signature_sequence
FT	Modified-site	483..497
FT	Modified-site	/label= Signature_sequence
FT	Modified-site	532..546
FT	Modified-site	/label= Signature_sequence
XX	WO958558-A2.	
XX	18-NOV-1999.	
XX	13-MAY-1999;	99WO-US10567.
XX	13-MAY-1998;	98US-0085343.
XX	26-AUG-1998;	98US-0098010.
XX	(INCY-)	INCYTE PHARM INC.
XX	Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C,	
XX	Baughn MR, Yang J;	
XX	WPI: 2000-086432/07.	
XX	N-PSDB: AA29233.	

PT Human cell signaling proteins useful for, e.g. diagnosing cell  
 PT proliferative and inflammatory disorders -  
 PS Claim 1; Page 77-78; 90pp; English.  
 CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded  
 CC by cDNA obtained from Incyte clone 3239149 of COLAUCT01 library. It is  
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is  
 CC found to be homologous to beta-transducin repeats containing  
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as  
 CC hybridisation probe for detecting CSIGP related sequences or allelic  
 CC variants. Recombinant CSIGP can be produced in host cells by transforming  
 CC them with genetically engineered vectors. Agonists or antagonists can be  
 CC used in the treatment of cell proliferative and inflammatory disorders  
 CC associated with decreased or increased CSIGP expression. CSIGP is used in  
 CC the diagnosis, prevention and treatment of cell proliferative disorders  
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.  
 SQ Sequence 569 AA;  
 Query Match 100.0%; Score 138; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EGHEELVRCIRPDNKRIVSGAYDGKI 26  
 DB 467 EGHEELVRCIRPDNKRIVSGAYDGKI 492  
 RESULT 14  
 AAB48298  
 ID AAB48298 standard; protein; 569 AA.  
 AC AAB48298;  
 XX  
 DT 02-APR-2001 (first entry)  
 DE Human ZF11 protein.  
 XX  
 XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;  
 KM CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
 KM Bad; Bcl-2; tumour; cytostatic.  
 XX  
 OS Homo sapiens.  
 PN WO200075184-A1.  
 PD 14-DEC-2000.  
 XX  
 PF 05-JUN-2000; 2000WO-US15449.  
 XX  
 PR 04-JUN-1999; 990S-0137494.  
 XX  
 PA (UYVA ) UNIV YALE.  
 PI Zhang H, Tsvetkov LM, Kondo T;  
 DR WPI; 2001-061703/07.  
 DR N-PSDB; AAC84610.  
 PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
 PT involves altering levels of proteins such as S-phase kinase associated  
 PT proteins 1, 2 and cullin/CDC53 proteins -  
 PS Claim 3; Page 130-132; 162pp; English.  
 CC The invention relates to methods of altering the polypeptide levels in a  
 CC cell, using proteins selected from S-phase kinase associated proteins 1  
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
 CC cullin/CDC53 family of proteins). The method is useful for altering the  
 CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for

CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
 CC that modulate interactions between Skp and target proteins are useful for  
 CC treating tumours.  
 SQ Sequence 569 AA;  
 Query Match 100.0%; Score 138; DB 22; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EGHEELVRCIRPDNKRIVSGAYDGKI 26  
 DB 467 EGHEELVRCIRPDNKRIVSGAYDGKI 492  
 RESULT 15  
 AAM78583  
 ID AAM78583 standard; protein; 579 AA.  
 AC AAM78583;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human protein SEQ ID NO 1245.  
 XX  
 DE Human protein SEQ ID NO 1245.  
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KM nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 PN WO200157190-A2.  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 DR WPI; 2001-476283/51.  
 DR N-PSDB; AAK51716.  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 PS Claim 20; Page 3504-3505; 6221pp; English.  
 CC The invention relates to polynucleotides (AAK51456-AAK51435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

CC (AAM60020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX  
 SQ Sequence 579 AA;

Query Match 100.0%; Score 138; DB 22; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-15;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHELVRCIRFDNKRIVSGAYDGI 26  
 ||||||||||||||||||  
 Db 477 EGHELVRCIRFDNKRIVSGAYDGI 502

Search completed: April 11, 2003, 11:48:17  
 Job time : 9.01218 secs

1000

1000



GenCore version 5.1.4\_p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 2.69184 Seconds  
(without alignments)  
284.191 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_467\_492

Perfect score: 138

Sequence: 1 EGHEILVRCIRFDNKRIVSGAYDGKI 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	517	1	US-08-190-802A-30
2	138	100.0	517	4	US-08-477-346-30
3	138	100.0	517	4	US-08-473-089-30
4	138	100.0	517	4	US-08-487-072A-30
5	133	96.4	29	1	US-08-190-802A-88
6	133	96.4	29	4	US-08-477-346-88
7	133	96.4	29	4	US-08-473-089-88
8	133	96.4	29	4	US-08-487-072A-88
9	76	55.1	269	3	US-08-899-578-6
10	74	53.6	387	3	US-08-899-578-2
11	63	45.7	318	1	US-08-190-802A-33
12	63	45.7	318	4	US-08-477-346-33
13	63	45.7	318	4	US-08-473-089-33
14	63	45.7	318	4	US-08-487-072A-33
15	62	44.9	640	1	US-08-177-165A-30
16	58	42.0	33	1	US-08-190-802A-103
17	58	42.0	33	4	US-08-477-346-103
18	58	42.0	33	4	US-08-473-089-103
19	58	42.0	33	4	US-08-487-072A-103
20	57	41.3	212	3	US-08-899-578-7
21	57	41.3	779	1	US-08-190-802A-32
22	57	41.3	779	4	US-08-477-346-32
23	57	41.3	779	4	US-08-473-089-32
24	57	41.3	779	4	US-08-487-072A-32
25	57	41.3	779	4	US-09-177-165A-29
26	56	40.6	210	4	US-09-291-170A-11
27	56	40.6	210	4	US-09-724-884-11

28	56	40.6	395	3	US-09-032-372-1	Sequence 1, Appl
29	55	39.9	29	1	US-08-190-802A-99	Sequence 99, Appl
30	55	39.9	29	4	US-08-477-346-99	Sequence 99, Appl
31	55	39.9	29	4	US-08-473-089-99	Sequence 99, Appl
32	55	39.9	29	4	US-08-487-072A-99	Sequence 99, Appl
33	55	39.9	1146	4	US-08-914-999-6	Sequence 6, Appl
34	52	37.7	29	1	US-08-190-802A-87	Sequence 87, Appl
35	52	37.7	29	4	US-08-477-346-87	Sequence 87, Appl
36	52	37.7	29	4	US-08-473-089-87	Sequence 87, Appl
37	52	37.7	29	4	US-08-487-072A-87	Sequence 87, Appl
38	52	37.7	606	2	US-08-883-534-3	Sequence 3, Appl
39	52	37.7	606	3	US-09-204-764-3	Sequence 3, Appl
40	52	37.7	906	1	US-08-190-802A-31	Sequence 31, Appl
41	52	37.7	906	4	US-08-477-346-31	Sequence 31, Appl
42	52	37.7	906	4	US-08-473-089-31	Sequence 31, Appl
43	52	37.7	906	4	US-08-487-072A-31	Sequence 31, Appl
44	51	37.0	34	1	US-08-190-802A-83	Sequence 83, Appl
45	51	37.0	34	4	US-08-477-346-83	Sequence 83, Appl

## ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
Sequence 30, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OR INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Delinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
MOLECULE TYPE: unknown  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-190-802A-30  
Query Match 100.0%; Score 138; DB 1; Length 517;  
Best Local Similarity 100.0%; Pred. No. 8.7e-14;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 431 EGHELVRCIRFDNKRIVSGAYDGKI 456

## RESULT 2

US-08-477-346-30  
; Sequence 30, Application US/08477346  
; Patent No. 6262023  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,346  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,072  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 517 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
; US-08-477-346-30

## Query Match

Best Local Similarity 100.0%; Score 138; DB 4; Length 517;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGHELVRCIRFDNKRIVSGAYDGKI 26

Db 431 EGHELVRCIRFDNKRIVSGAYDGKI 456

## RESULT 3

US-08-473-089-30  
; Sequence 30, Application US/08473089  
; Patent No. 6342368  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-473-089-30

## Query Match

Best Local Similarity 100.0%; Score 138; DB 4; Length 517;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGHELVRCIRFDNKRIVSGAYDGKI 26

Db 431 EGHELVRCIRFDNKRIVSGAYDGKI 456

## RESULT 4

US-08-487-072A-30  
; Sequence 30, Application US/08487072A  
; Patent No. 6423684  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,072A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 INFORMATION FOR SEQ ID NO: 30:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 517 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
 US-08-487-072A-30

Query Match 100.0%; Score 138; DB 4; Length 517;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-14;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHELVRCIRFDNKRIVSGAYDGKI 26  
 Db 431 EGHELVRCIRFDNKRIVSGAYDGKI 456

RESULT 5  
 US-08-190-802A-88  
 Sequence 88, Application US/08190802A  
 Patent No. 5519003  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thee of  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Delinger & Associates  
 STREET: P.O. Box 60850  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94306-0850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/190.802A  
 FILING DATE: 01-FEB-1994  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fabian, Gary R.  
 REGISTRATION NUMBER: 33,875  
 REFERENCE/DOCKET NUMBER: 8600-0139  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 88:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 29 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: BETA TRCP rVI, Fig. 13  
 US-08-190-802A-88

Query Match 96.4%; Score 133; DB 1; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYDGKI 26  
 Db 2 GHEELVRCIRFDNKRIVSGAYDGKI 26

Db 1 GHEELVRCIRFDNKRIVSGAYDGKI 25

RESULT 6  
 US-08-477-346-88  
 Sequence 88, Application US/08477346  
 Patent No. 6262023  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thee of  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 2000 Pennsylvania Avenue, NW  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20006-1812  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/477,346  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/487,072  
 FILING DATE: 07-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MURASHIGE, KATE H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 2550-0025.20  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 INFORMATION FOR SEQ ID NO: 88:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 29 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: BETA TRCP rVI, Fig. 13  
 US-08-477-346-88

Query Match 96.4%; Score 133; DB 4; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYDGKI 26  
 Db 1 GHEELVRCIRFDNKRIVSGAYDGKI 25

RESULT 7  
 US-08-473-089-88  
 Sequence 88, Application US/08473089  
 Patent No. 6342368  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thee of  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVI, Fig. 13  
US-08-473-089-88

Query Match  
Best Local Similarity 96.4%; Score 133; DB 4; Length 29;  
Matches 25; Conservativity 100.0%; Pred. No. 1,9e-14;  
Mismatches 0; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYDGKI 26  
DB 1 GHEELVRCIRFDNKRIVSGAYDGKI 25

RESULT 8  
US-08-487-072A-88  
Sequence 88, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theoreof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVI, Fig. 13  
US-08-487-072A-88

Query Match  
Best Local Similarity 96.4%; Score 133; DB 4; Length 29;  
Matches 25; Conservativity 100.0%; Pred. No. 1,9e-14;  
Mismatches 0; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYDGKI 26  
DB 1 GHEELVRCIRFDNKRIVSGAYDGKI 25

RESULT 9  
US-08-899-578-6  
Sequence 6, Application US/08899578  
Patent No. 6087153  
GENERAL INFORMATION:  
APPLICANT: Greenwald, Iva  
APPLICANT: Hubbard, E. Jane  
TITLE OF INVENTION: SEL-10 AND USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,578  
FILING DATE: 24-JUL-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AKC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 278-0525  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 209 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-899-578-6

Query Match  
Best Local Similarity 55.1%; Score 76; DB 3; Length 209;  
Matches 12; Conservativity 52.2%; Pred. No. 0.00018;  
Mismatches 7; Indels 0; Gaps 0;

QY 1 EGHEELVRCIRFDNKRIVSGAYD 23  
DB 90 DGHAAVRCVQFDGTIVSGGYD 112

RESULT 10  
US-08-899-578-2  
Sequence 2, Application US/08899578  
Patent No. 6087153  
GENERAL INFORMATION:  
APPLICANT: Greenwald, Iva  
APPLICANT: Hubbard, E. Jane  
TITLE OF INVENTION: SEL-10 AND USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,578  
FILING DATE: 24-JUL-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/53200/3PM/AMC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 278-0525  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 587 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-899-578-2

Query Match 53.6%; Score 74; DB 3; Length 587;  
Best Local Similarity 54.5%; Pred. No. 0.0012;  
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 CHEELVRCIRPDNKRIVSGAYD 23  
||| ||| : ||| ||| |||  
DB 376 GHAAVRCVDPDGTTVSGGSD 397

RESULT 11  
US-08-190-802A-33  
Sequence 33, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A

FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 amino acids  
TYPE: amino acid  
MOLECULE TYPE: unknown  
TOPOLOGY: unknown  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: GBLP -CHLAMIDOMONAS HOMOLOG, Fig. 16  
US-08-190-802A-33

Query Match 45.7%; Score 63; DB 1; Length 318;  
Best Local Similarity 51.9%; Pred. No. 0.032;  
Matches 14; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

OY 1 EGHEELVRCIRPD---NKRIVSGAYD 23  
||| | | : ||| : |  
DB 147 EGHTEWVSCVRFSPMTNPIVSGGWD 173

RESULT 12  
US-08-477-346-33  
Sequence 33, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO

ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: GBLP -CHLAMYDOMONAS HOMOLOG, Fig. 16  
US-08-473-089-33

Query Match 45.7%; Score 63; DB 4; Length 318;  
Best Local Similarity 51.9%; Pred. No. 0.032;  
Matches 14; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 1 EGHEELVRCIRFD---NKRIVSGAYD 23  
DB 147 EGHEWVSCVRFSPMTNPIIVSGWMD 173

RESULT 13  
US-08-473-089-33  
Sequence 33, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morlison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: GBLP -CHLAMYDOMONAS HOMOLOG, Fig. 16  
US-08-473-089-33

Query Match 45.7%; Score 63; DB 4; Length 318;  
Best Local Similarity 51.9%; Pred. No. 0.032;  
Matches 14; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 1 EGHEELVRCIRFD---NKRIVSGAYD 23  
DB 147 EGHEWVSCVRFSPMTNPIIVSGWMD 173

RESULT 14  
US-08-487-072A-33  
Sequence 33, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morlison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: GBLP -CHLAMYDOMONAS HOMOLOG, Fig. 16  
US-08-487-072A-33

Query Match 45.7%; Score 63; DB 4; Length 318;  
Best Local Similarity 51.9%; Pred. No. 0.032;  
Matches 14; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 1 EGHEELVRCIRFD---NKRIVSGAYD 23  
DB 147 EGHEWVSCVRFSPMTNPIIVSGWMD 173

RESULT 15  
US-09-177-165A-30  
Sequence 30, Application US/09177165A  
Patent No. 6426205  
GENERAL INFORMATION:

APPLICANT: Tyters, Mike  
APPLICANT: Williams, Andrew  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBQUITIN  
TITLE OF INVENTION: DEPENDENT PROTEOLYSIS  
FILE REFERENCE: 11757.10USU1  
CURRENT APPLICATION NUMBER: US/09/177,165A  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: 60/092,443  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/063,254  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 30  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-177-165A-30

Query Match 44.98; Score 62; DB 4; Length 640;  
 Best Local Similarity 56.08; Pred. No. 0.1;  
 Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 Oy 2 GHELVRCIRFDNKRIVSGAYDGKI 26  
 ||||| ||||:||||  
 Db 550 GHVGVMDIADNFRITSSHDGSI 574

Search completed: April 10, 2003, 12:51:19  
 Job time : 2.69184 secs

1. The first part of the paper is devoted to a discussion of the general principles of the theory of the structure of the human brain, and to a description of the various methods of investigation which have been employed in the study of the brain.



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:49:06 : Search time 2.66017 Seconds

(without alignments)  
597.532 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_467\_492  
Perfect score: 138  
Sequence: 1 ECHHELVCIRFDNKRIVSGAYDGI 26

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications-AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCIT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCITUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	100.0	569	9 US-10-038-010-8	Sequence 8, Appl1
2	138	100.0	569	12 US-10-042-417-2	Sequence 2, Appl1
3	82	59.4	540	9 US-09-213-888-7	Sequence 7, Appl1
4	82	59.4	540	9 US-09-213-888-10	Sequence 10, Appl1
5	82	59.4	540	9 US-09-328-877A-7	Sequence 7, Appl1
6	82	59.4	540	9 US-09-328-877A-10	Sequence 10, Appl1
7	82	59.4	545	9 US-09-213-888-6	Sequence 6, Appl1
8	82	59.4	545	9 US-09-328-877A-6	Sequence 6, Appl1
9	82	59.4	553	9 US-09-213-888-5	Sequence 5, Appl1
10	82	59.4	553	9 US-09-328-877A-5	Sequence 5, Appl1
11	82	59.4	559	9 US-09-213-888-9	Sequence 9, Appl1
12	82	59.4	559	9 US-09-328-877A-9	Sequence 9, Appl1
13	82	59.4	589	9 US-09-213-888-8	Sequence 8, Appl1
14	82	59.4	589	9 US-09-328-877A-8	Sequence 8, Appl1
15	82	59.4	592	9 US-09-213-888-4	Sequence 4, Appl1
16	82	59.4	592	9 US-09-328-877A-4	Sequence 4, Appl1
17	82	59.4	626	9 US-09-213-888-21	Sequence 21, Appl1
18	82	59.4	626	9 US-09-328-877A-21	Sequence 21, Appl1
19	82	59.4	627	9 US-09-213-888-3	Sequence 3, Appl1

20	82	59.4	627	9 US-09-328-877A-3	Sequence 3, Appl1
21	82	59.4	666	9 US-09-213-888-27	Sequence 27, Appl1
22	82	59.4	666	9 US-09-328-877A-27	Sequence 27, Appl1
23	82	59.4	669	9 US-09-213-888-25	Sequence 25, Appl1
24	82	59.4	669	9 US-09-328-877A-25	Sequence 25, Appl1
25	71	51.4	678	10 US-09-801-368-314	Sequence 314, App
26	62	44.9	640	9 US-10-060-019-30	Sequence 30, Appl1
27	57	41.3	779	9 US-10-060-019-29	Sequence 29, Appl1
28	56	40.6	353	10 US-09-326-248B-6	Sequence 6, Appl1
29	56	40.6	371	10 US-09-925-297-699	Sequence 699, App
30	55	39.9	1146	9 US-09-832-292-10	Sequence 10, Appl1
31	55	39.9	1146	10 US-09-994-485-6	Sequence 6, Appl1
32	54	39.1	1198	10 US-09-866-582-36	Sequence 36, Appl1
33	54	39.1	1241	12 US-10-001-215-5	Sequence 5, Appl1
34	54	39.1	1356	9 US-10-077-111-10	Sequence 10, Appl1
35	53.5	38.8	392	10 US-09-925-300-1085	Sequence 1085, Ap
36	53	38.4	464	10 US-09-925-302-775	Sequence 775, App
37	52	37.7	744	10 US-09-925-300-1347	Sequence 1347, Ap
38	50.5	36.6	1194	10 US-09-876-667-2	Sequence 2, Appl1
39	50.5	36.6	1205	10 US-09-876-667-16	Sequence 16, Appl1
40	50	36.2	114	9 US-09-796-692-1541	Sequence 1541, Ap
41	49	35.5	44	10 US-09-843-845-26	Sequence 26, Appl1
42	49	35.5	70	9 US-09-796-692-1822	Sequence 1822, Ap
43	49	35.5	70	9 US-09-796-692-2077	Sequence 2077, Ap
44	49	35.5	114	9 US-09-796-692-1445	Sequence 1445, Ap
45	49	35.5	114	9 US-09-796-692-1895	Sequence 1895, Ap

## ALIGNMENTS

RESULT 1  
US-10-038-010-8  
Sequence 8, Application US/10038010  
Publication No. US20030040089A1  
GENERAL INFORMATION:  
APPLICANT: Pierre, Legrain  
TITLE OF INVENTION: Protein-protein interactions in adipocyte cells  
FILE REFERENCE: B4767A  
CURRENT FILING DATE: 2002-07-23  
PRIOR APPLICATION NUMBER: US/60/259,377  
PRIOR FILING DATE: 2001-01-02  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: beta-TCP1  
LOCATION: (1)..(569)  
OTHER INFORMATION:  
US-10-038-010-8

Query Match 100.0%; Score 138; DB 9; Length 569;  
Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 26; Conservative 0; Mismatches 0; Indels 0;

Qy 1 ECHHELVCIRFDNKRIVSGAYDGI 26  
Db 467 ECHHELVCIRFDNKRIVSGAYDGI 492

RESULT 2  
US-10-042-417-2  
Sequence 2, Application US/10042417  
Patent No. US20020123082A1  
GENERAL INFORMATION:  
APPLICANT: Pagano, M.  
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS

```
FILE REFERENCE: 5914-090-999
CURRENT APPLICATION NUMBER: US/10/042,417
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: 60/250,179
PRIOR FILING DATE: 2001-01-5
NUMBER OF SEQ ID NOS: 89
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 569
TYPE: PRT
ORGANISM: Homo sapiens
US-10-042-417-2

Query Match          100.0%; Score 138; DB 12; Length 569;
Best Local Similarity 100.0%; Pred. No. 2,6e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHELVRCIRFDNKRIVSGAYDGI 26
DB 467 EGHELVRCIRFDNKRIVSGAYDGI 492

RESULT 3
US-09-213-888-7
Sequence 7, Application US/09213888A
Patent No. US20020164683A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Li, Jinhe
APPLICANT: Pauley, Adele M.
APPLICANT: Pharmacia & Upjohn Company
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
FILE REFERENCE: 6142
CURRENT APPLICATION NUMBER: US/09/213,888A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 7
LENGTH: 540
TYPE: PRT
ORGANISM: Homo sapiens
US-09-213-888-7

Query Match          59.4%; Score 82; DB 9; Length 540;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23
DB 332 GHVAAVRCVQYDGRRVVSGAYD 353

RESULT 4
US-09-213-888-10
Sequence 10, Application US/09213888A
Patent No. US20020164683A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Li, Jinhe
APPLICANT: Pauley, Adele M.
APPLICANT: Pharmacia & Upjohn Company
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
FILE REFERENCE: 6142
CURRENT APPLICATION NUMBER: US/09/213,888A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 10
LENGTH: 540
TYPE: PRT
ORGANISM: Homo sapiens
```

```
US-09-213-888-10

Query Match          59.4%; Score 82; DB 9; Length 540;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23
DB 332 GHVAAVRCVQYDGRRVVSGAYD 353

RESULT 5
US-09-328-877A-7
Sequence 7, Application US/09328877A
Patent No. US20020177187A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Li, Jinhe
APPLICANT: Pauley, Adele M.
APPLICANT: Pharmacia & Upjohn Company
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
FILE REFERENCE: 6142
CURRENT APPLICATION NUMBER: US/09/328,877A
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 7
LENGTH: 540
TYPE: PRT
ORGANISM: Homo sapiens
US-09-328-877A-7

Query Match          59.4%; Score 82; DB 9; Length 540;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23
DB 332 GHVAAVRCVQYDGRRVVSGAYD 353

RESULT 6
US-09-328-877A-10
Sequence 10, Application US/09328877A
Patent No. US20020177187A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Li, Jinhe
APPLICANT: Pauley, Adele M.
APPLICANT: Pharmacia & Upjohn Company
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
FILE REFERENCE: 6142
CURRENT APPLICATION NUMBER: US/09/328,877A
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 10
LENGTH: 540
TYPE: PRT
ORGANISM: Homo sapiens
US-09-328-877A-10

Query Match          59.4%; Score 82; DB 9; Length 540;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23
DB 332 GHVAAVRCVQYDGRRVVSGAYD 353
```

```
RESULT 7
US-09-213-888-6
; Sequence 6, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-6

Query Match          59.4%; Score 82; DB 9; Length 545;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 GHEELVRCIRFDNKRIVSGAYD 23
  || |||:::|:|:|||||
Db 337 GHVAAVRCVQYDGRRVSGAYD 358

RESULT 8
US-09-328-877A-6
; Sequence 6, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-6

Query Match          59.4%; Score 82; DB 9; Length 545;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 GHEELVRCIRFDNKRIVSGAYD 23
  || |||:::|:|:|||||
Db 337 GHVAAVRCVQYDGRRVSGAYD 358

RESULT 9
US-09-213-888-5
; Sequence 5, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-5

Query Match          59.4%; Score 82; DB 9; Length 553;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 GHEELVRCIRFDNKRIVSGAYD 23
  || |||:::|:|:|||||
Db 345 GHVAAVRCVQYDGRRVSGAYD 366
```

```
FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-5

Query Match          59.4%; Score 82; DB 9; Length 553;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 GHEELVRCIRFDNKRIVSGAYD 23
  || |||:::~:~:|||||
Db 345 GHVAAVRCVQYDGRRVSGAYD 366

RESULT 10
US-09-328-877A-5
; Sequence 5, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-5

Query Match          59.4%; Score 82; DB 9; Length 553;
Best Local Similarity 59.1%; Pred. No. 0.0001;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 GHEELVRCIRFDNKRIVSGAYD 23
  || |||:::~:~:|||||
Db 345 GHVAAVRCVQYDGRRVSGAYD 366

RESULT 11
US-09-213-888-9
; Sequence 9, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 9
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-9
```

Search completed: April 10, 2003, 13:16:47  
Job time : 3.66017 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 3.00853 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_467\_492  
Perfect score: 138  
Sequence: 1 EGHELVRCIRFDNKRIVSGAYDGKI 26

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR.73:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	100.0	518	2	B48088
2	132	95.7	701	2	T16607
3	92	66.7	267	2	S62507
4	92	66.7	506	2	T50211
5	74	53.6	579	2	T22703
6	68	49.3	1101	2	T26919
7	63	45.7	318	2	S11904
8	63	45.7	605	2	T38932
9	62	44.9	601	2	TJ0238
10	62	44.9	640	2	S49932
11	61	44.2	615	2	S54468
12	59	42.8	650	2	T46660
13	59	42.8	659	2	S38108
14	59	42.8	1888	2	T14273
15	58	42.0	357	2	AT2099
16	57.5	41.7	386	2	T26805
17	57	41.3	440	2	T01768
18	57	41.3	522	2	D96764
19	57	41.3	580	2	G72202
20	57	41.3	779	2	S56245
21	56.5	40.9	494	2	T40471
22	56.5	40.9	586	2	T19756
23	56	40.6	283	2	T17256
24	56	40.6	325	2	G87344
25	56	40.6	353	2	A54969
26	56	40.6	714	2	S56893
27	55	39.9	1146	2	A55532
28	54	39.1	334	2	T03764
29	54	39.1	586	2	T38992

30	54	39.1	651	2	T50289	WD repeat protein
31	54	39.1	1198	2	T43484	hypothetical prote
32	54	39.1	1356	2	T18521	beta transducin-11
33	53	38.4	304	2	A61837	WD-40 repeat prote
34	53	38.4	314	2	T43299	probable GTP-bind
35	53	38.4	316	2	S57839	CPC2 protein - Neu
36	53	38.4	438	2	T45823	hypothetical prote
37	53	38.4	510	2	T20276	hypothetical prote
38	53	38.4	745	2	S74281	hypothetical prote
39	52.5	38.0	1043	2	T23875	hypothetical prote
40	52	37.7	370	2	A66741	hypothetical wd-40
41	52	37.7	399	2	T38388	WDRI protein - hum
42	52	37.7	606	2	T13152	Colg1-associated p
43	52	37.7	906	2	S35342	coatomer complex b
44	52	37.7	906	2	S35312	hypothetical prote
45	51	37.0	395	2	T23317	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

B48088  
beta-transducin repeat-containing protein - African clawed frog  
N:Alternate names: beta-Trop  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000  
C:Accession: B48088  
R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.  
Mol. Cell. Biol. 13, 4953-4966, 1993  
A:Title: Saccharomyces cerevisiae CDC15 mutants arrested at a late stage in anaphase  
A:Reference number: A48088; M0ID:93330289; PMID:8393141  
A:Accession: B48088  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-518 <SPE>  
A:Cross-References: GB:98268; NID:9295542; PIDN:AAA02810.1; PID:9295543  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: duplication  
F:431-462/Domain: WD repeat homology <WD1>

Query Match 100.0%; Score 138; DB 2; Length 518;  
Best Local Similarity 100.0%; Pred. No. 6e-13;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EGHELVRCIRFDNKRIVSGAYDGKI 26  
DB 432 EGHELVRCIRFDNKRIVSGAYDGKI 457

##### RESULT 2

T16607  
hypothetical protein K10B2.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16607  
R:Miller, N.  
Submitted to the EMBL Data Library, June 1995  
A:Description: The sequence of C. elegans cosmid K10B2.  
A:Reference number: 218545  
A:Accession: T16607  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-701 <MTL>  
A:Cross-References: EMBL:U28730; NID:9860694; PID:9860695; PIDN:AAA68258.1; CESP:K10B  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:K10B2.1  
A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3  
Query Match 95.7%; Score 132; DB 2; Length 701;  
Best Local Similarity 96.2%; Pred. No. 6.8e-12;  
Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EGHLELVRCIRFDNKRIVSGAYDGKI 26  
 |||||  
 Db 458 EGHLELVRCIRFDNKRIVSGAYDGKI 483

## RESULT 3

hypothetical trp-asp repeat-containing protein - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: T38502; S62507  
 R:Jones, L.; Murphy, L.; McNeil, A.; Simpson, I.; Harris, D.; Barrell, B.G.; Rajandream, A.  
 submitted to the EMBL Data Library, October 1995  
 A:Reference number: T21798  
 A:Accession: T38502  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-267 <T02>  
 A:Cross-references: EMBL:Z66525; NID:g1044926; PIDN:CA91423.1; PID:g1044927; GSPDB:GNOC  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 F:18-51/Domain: WD repeat homology <WD1>  
 F:104-137/Domain: WD repeat homology <WD2>

Query Match 66.7%; Score 92; DB 2; Length 267;  
 Best Local Similarity 53.8%; Pred. No. 2.9e-06;  
 Matches 14; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 1 EGHLELVRCIRFDNKRIVSGAYDGKI 26  
 |||||  
 Db 187 QGHEDLIRTVRFNDEKIVSGYDGTV 212

## RESULT 4

WD-repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000  
 C:Accession: T50211  
 R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.; Mc  
 submitted to the EMBL Data Library, January 2000  
 A:Reference number: Z25046  
 A:Accession: T50211  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-506 <MC0D>  
 A:Cross-references: EMBL:AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05  
 A:Experimental source: strain 972h(-); cosmid c30  
 C:Genetics:  
 A:Gene: SPAC29E6.01; SPDB:SPAC30.05  
 A:Map position: 1  
 A:Introns: 43/1; 74/3  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 66.7%; Score 92; DB 2; Length 506;  
 Best Local Similarity 53.8%; Pred. No. 5.7e-06;  
 Matches 14; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 1 EGHLELVRCIRFDNKRIVSGAYDGKI 26  
 |||||  
 Db 426 QGHEDLIRTVRFNDEKIVSGYDGTV 451

RESULT 5  
 T22703  
 hypothetical protein F55B12.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T22703  
 R:Sims, M.  
 submitted to the EMBL Data Library, September 1996

A:Reference number: Z19602  
 A:Accession: T22703  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-579 <WTL>  
 A:Cross-references: EMBL:Z79757; PIDN:CAB02129.1; GSPDB:GN00023; CESP:F55B12.3  
 A:Experimental source: clone F55B12  
 C:Genetics:  
 A:Gene: CESP:F55B12.3  
 A:Map position: 5  
 A:Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

Query Match 53.6%; Score 74; DB 2; Length 579;  
 Best Local Similarity 54.5%; Pred. No. 0.0036;  
 Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 GHEELVRCIRFDNKRIVSGAYD 23  
 |||||  
 Db 368 GHAAVRCVQFDGTVYSSGVD 389

## RESULT 6

hypothetical protein Y45F10B.10 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T26919  
 R:McMurray, A.  
 submitted to the EMBL Data Library, January 1998  
 A:Reference number: Z20286  
 A:Accession: T26919  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1101 <WTL>  
 A:Cross-references: EMBL:AL021487; PIDN:CA16357.1; GSPDB:GN00022; CESP:Y45F10B.10  
 A:Experimental source: clone Y45F10B  
 C:Genetics:  
 A:Gene: CESP:Y45F10B.10  
 A:Map position: 4  
 A:Introns: 217/3; 292/1; 451/3; 556/2; 637/3; 755/3; 823/2; 946/1; 1078/3

Query Match 49.3%; Score 68; DB 2; Length 1101;  
 Best Local Similarity 55.6%; Pred. No. 0.058;  
 Matches 15; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

OY 2 GHEELVRCIRFDNKRIVSGAYDGKI 26  
 |||||  
 Db 885 GHENVTCCISFDERIVSGARDEKI 911

## RESULT 7

GTP-binding regulatory protein beta chain homolog - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 21-Jul-2000  
 C:Accession: S11904  
 R:Schloss, J.A.  
 M.O. Gen. Genet. 221, 443-452, 1990  
 A:Title: A Chlamydomonas gene encodes a G protein beta subunit-like polypeptide.  
 A:Reference number: S11904; MUID:90340296; PMID:2116589  
 A:Accession: S11904  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-318 <MOL>  
 A:Cross-references: EMBL:X53574; NID:g18144; PIDN:CA37638.1; PID:g18145  
 C:Genetics:  
 A:Introns: 37/2; 120/3; 177/3; 265/2  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 F:11-46/Domain: WD repeat homology <WD1>  
 F:60-93/Domain: WD repeat homology <WD2>  
 F:102-135/Domain: WD repeat homology <WD3>  
 F:146-181/Domain: WD repeat homology <WD4>  
 F:190-223/Domain: WD repeat homology <WD5>







GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 12:51:26 ; Search time 1.86845 seconds  
(without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_467\_492  
Perfect score: 138  
Sequence: 1 EGHEELVRCIRFDNKRIVSGAYDKRI 26

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	100.0	518	TRCB_XENLA	091854 xenopus lae
2	138	100.0	542	FWIB_HUMAN	094kb1 homo sapien
3	138	100.0	605	FWIA_HUMAN	094297 homo sapien
4	132	99.7	665	L123_CAEL	009990 caenorhabdi
5	92	66.7	506	POFB_SCHPO	009855 schizosacch
6	74	53.6	579	SEIO_CAEL	093794 caenorhabdi
7	71	51.4	678	SCOB_EMEI	000659 caenorhabdi
8	66	47.8	1261	APAF_BRARE	091988 brachydanio
9	63	45.7	318	GRUP_CHURE	P2387 chlamydomon
10	63	45.7	605	POPI_SCHPO	P87053 schizosacch
11	62	44.9	601	WD66_PHYRO	P90587 physarum po
12	62	44.9	640	MT30_YEAST	P39014 saccharomyc
13	61	44.2	615	ATPI_YEAST	P46680 saccharomyc
14	59	42.8	650	SCO2_NEUCR	001277 neurospora
15	59	42.8	659	YK16_YEAST	P36130 saccharomyc
16	57	41.3	779	CC4_YEAST	P07834 saccharomyc
17	56	40.6	353	GBB5_HUMAN	014775 homo sapien
18	56	40.6	353	GBB5_MOUSE	P54314 mus musculu
19	56	40.6	714	Y1L2_YEAST	P47025 saccharomyc
20	55	39.9	1146	KMHA_DICDI	P45279 dictyosteli
21	54	39.1	317	GLRP_OREMI	042249 oreochromis
22	54	39.1	334	GLRP_ORYSA	P49027 oryza sativ
23	54	39.1	586	T012_SCHPO	094u98 schizosacch
24	54	39.1	651	Y1X1_SCHPO	094u98 schizosacch
25	54	39.1	1242	WDRA_HUMAN	094b96 homo sapien
26	54	39.1	1356	HEM1_PODAN	000808 podospora a
27	53	38.4	314	GLRP_SCHPO	010281 schizosacch
28	53	38.4	316	GLRP_NEUCR	001369 neurospora
29	53	38.4	608	W0R1_XENLA	094712 xenopus lae
30	53	38.4	682	Y124_HUMAN	014137 homo sapien
31	53	38.4	759	YCD9_YEAST	P25569 saccharomyc
32	52	37.7	318	GLRP_DROME	018640 drosophilla
33	52	37.7	327	GLRP_ARATH	024456 arabidopsis

## ALIGNMENTS

```

RESULT 1
ID      TRCB_XENLA      STANDARD:      PRT:      518 AA.
AC      091854: P70037; P70038:
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Beta-TRCP (Beta-transducin repeat-containing protein).
GN      FBXW1 OR BTBRCP.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC      Xenopodidae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      MEDLINE=9330289; PubMed=8393141;
RX      MEDLINE=97109804; PubMed=8952061;
RA      Spevak W., Keiper B.D., Stratowa C., Gastanon M.J.;
RT      "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
RT      anaphase are rescued by Xenopus CDNAs encoding N-ras or a protein
RT      with beta-transducin repeats."
RL      Mol. Cell. Biol. 13:4953-4966(1993).
RN      [2]
RP      SEQUENCE OF 302-518 FROM N.A.
RX      MEDLINE=97109804; PubMed=8952061;
RA      Hudson J.W., Alarcon V.B., Elinson R.P.;
RT      "Identification of new localized RNAs in the Xenopus oocyte by
RT      differential display PCR."
RL      Dev. Genet. 19:190-198(1996).
CC      -1- FUNCTION: Probably recognizes and binds to some phosphorylated
CC      proteins and promotes their ubiquitination and degradation.
CC      -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX
CC      (BY SIMILARITY).
CC      -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
CC      MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
CC      GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
CC      NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
CC      TADPOLE EMBRYO.
CC      -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC      -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC      -1- SIMILARITY: STRONG. TO C.ELEGANS K10B2.1.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL collaboration -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M98268; AAA02810.1; -
DR      EMBL; U63921; AAB4672.1; -
DR      EMBL; U63922; AAB4672.1; -
DR      InterPro: IPR001810; F-box.
DR      InterPro: IPR001680; WD40.
DR      Pfam: PF00400; WD40; 7.

```

DR Pfam: PR00646; F-box; 1.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ubl conjugation pathway; Repeat; WD repeat.  
 KW DOMAIN 119 157 F-BOX.  
 FT REPEAT 230 258 WD 1.  
 FT REPEAT 270 298 WD 2.  
 FT REPEAT 310 338 WD 3.  
 FT REPEAT 353 381 WD 4.  
 FT REPEAT 393 421 WD 5.  
 FT REPEAT 433 461 WD 6.  
 FT REPEAT 482 510 WD 7.  
 FT REPEAT 516 548 GEM -> EFR (IN REF. 2).  
 FT CONFLICT 302 304 GEM -> AAH (IN REF. 2).  
 FT CONFLICT 516 518  
 SQ SEQUENCE 518 AA: 59507 MM: 2A52EC19028127F3 CRC64;

Query Match 100.0%; Score 138; DB 1; Length 518;  
 Best local similarity 100.0%; Pred. No. 1.5e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EGHLELYRCIRFDNKRIVSGAYDGKI 26  
 Db 432 EGHLELYRCIRFDNKRIVSGAYDGKI 457

## RESULT 2

FW1A\_HUMAN STANDARD; PRT; 542 AA.

AC Q9Y297; Q9Y297; Q9P2S8; Q9P2S9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TTCP2).  
 GN FBXW1B OR FBW1B OR BTCP2 OR KIAA0696.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Cenciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RC TISSUE-Petal lung;  
 RX MEDLINE=20160456; PubMed=10694485;  
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,  
 RA Katoh M.;  
 RT "Molecular cloning and genomic structure of the betaTRCP2 gene on  
 RT chromosome 5q35.1";  
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kocani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A, B AND C (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF176022; AAF04528.1; -  
 DR EMBL: AB033279; BAA92329.1; -  
 DR EMBL: AB033280; BAA92330.1; -  
 DR EMBL: AB033281; BAA92331.1; -  
 DR EMBL: AB014596; BAA1671.1; ALT\_INIT.  
 DR Genew; HGNC:13607; FBXW1B.  
 DR MIM: 605651; -  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR Pfam; PF00646; F-box; 1.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 KW DOMAIN 129 167 F-BOX.  
 FT REPEAT 238 275 WD 1.  
 FT REPEAT 278 315 WD 2.  
 FT REPEAT 318 355 WD 3.  
 FT REPEAT 361 398 WD 4.  
 FT REPEAT 401 440 WD 5.  
 FT REPEAT 442 478 WD 6.  
 FT REPEAT 490 527 WD 7.  
 FT REPEAT 527 542 MISSING (IN ISOFORM A).  
 FT VARSPLIC 16 49 CSVPRLMIGCANVESKATCSQSPSVRL -> NTSV  
 FT VARSPLIC 16 48 MEDONEDSEPKNTLM (IN ISOFORM B).  
 SQ SEQUENCE 542 AA: 62090 MM: 7CD40087EFA5C8A CRC64;

Query Match 100.0%; Score 138; DB 1; Length 542;  
 Best local similarity 100.0%; Pred. No. 1.6e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EGHLELYRCIRFDNKRIVSGAYDGKI 26  
 Db 440 EGHLELYRCIRFDNKRIVSGAYDGKI 465

## RESULT 3

FW1A\_HUMAN STANDARD; PRT; 605 AA.

AC Q9Y297; Q9Y213;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TTCP)  
 DE (E3S1Kappab) (pikappaBalpha-E3 receptor subunit).  
 GN FBXW1A OR FBW1A OR BTCP OR BTCP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=99075339; PubMed=9859996;

RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,  
 RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
 RT "Identification of the receptor component of the Ikapalpha-  
 RT ubiquitin ligase.";  
 RL Nature 396:590-594(1998).  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Lymphoid;  
 RX MEDLINE=98325370; PubMed=9660940;  
 RA Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,  
 RA Thomas D., Strebel K., Benarous R.;  
 RT "A novel human WD protein, h-beta TRCP, that interacts with HIV-1 Vpu  
 RT connects CD4 to the ER degradation pathway through an F-box motif.";  
 RL Mol. Cell 1:565-574(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Cenciarelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99145464; PubMed=9990852;  
 RA Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,  
 RA Harper J.W.;  
 RT "The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically  
 RT with phosphorylated destruction motifs in I-kappa-B-alpha and  
 RT beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro.";  
 RL Genes Dev. 13:270-283(1999).  
 CC -1- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA  
 CC (PITAPPAALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR  
 CC UBIQUITINATION AND DEGRADATION.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF101784; AAC08702.1;  
 DR EMBL: Y14153; CAAT74572.1;  
 DR EMBL: AF129530; AAF04464.1;  
 DR Genew: HGNC:1144; BTRC.  
 DR MIM: 603482;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF004400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ub1 conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 DR DOMAIN 190 228 F-BOX.  
 FT REPEAT 301 338 WD 1.  
 FT REPEAT 341 378 WD 2.  
 FT REPEAT 381 418 WD 3.  
 FT REPEAT 424 461 WD 4.  
 FT REPEAT 464 503 WD 5.  
 FT REPEAT 505 541 WD 6.

FT REPEAT 553 590 WD 7.  
 FT VARSPLIC 17 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 605 AA; 68866 MW; 4C673B7E400FD37 CRC64;  
 Query Match 100.0%; Score 138; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 1,8e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ECHELVRCIRPDNKRIVSGAVDGKI 26  
 Db 503 ECHELVRCIRPDNKRIVSGAVDGKI 528  
 RESULT 4  
 L123\_CAEEL STANDARD: PRT: 665 AA.  
 AC Q09990; Q9GNM6;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein 1ln-23.  
 GN 1ln-23 OR K10B2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION: DEVELOPMENTAL STAGE, AND MUTAGENESIS.  
 RX MEDLINE=20515608; PubMed=11060233;  
 RA Kipreos E.T., Gonel S.P., Hedgecock E.M.;  
 RT "The Caenorhabditis elegans F-box/WD-repeat protein 1ln-23 functions  
 RT to limit cell division during development.";  
 RL Development 127:5071-5082(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol NZ;  
 RA Miller N.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Functions cell autonomously to negatively regulate cell  
 CC cycle progression. Required to restrain cell proliferation in  
 CC response to developmental cues. Probably recognizes and binds to  
 CC some proteins and promotes their ubiquitination and degradation  
 CC (By similarity).  
 CC -1- SUBUNIT: Part of a SCF (SKP1-CULLIN-F-box) protein ligase complex  
 CC (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- DEVELOPMENTAL STAGE: Highest levels in embryos and adults, lowest  
 CC levels in larvae. Maternal expression results in high zygotic  
 CC levels.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO X.LAEVIS PBXW1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF275253; AAG28037.1;  
 DR EMBL: U28730; AAA68258.2;  
 DR WormPep: K10B2.1; CE28600.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF004400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.

```

DR ProDom: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Developmental protein; Cell cycle; Cell division;
KW Ubl conjugation pathway; Repeat; WD repeat.
FT DOMAIN 81 127 F-BOX.
FT REPEAT 220 257 WD 1.
FT REPEAT 260 299 WD 2.
FT REPEAT 301 337 WD 3.
FT REPEAT 343 380 WD 4.
FT REPEAT 383 420 WD 5.
FT REPEAT 423 460 WD 6.
FT REPEAT 472 509 WD 7.
FT DOMAIN 525 582 ALA-RICH.
FT MUTAGEN 441 441 G->R: IN LIN-32(RH293).
SO SEQUENCE 665 AA; 75916 MW; BF3E9AF51F12ECCC CRC64;

Query Match
Best Local Similarity 95.7%; Score 132; DB 1; Length 665;
Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGHEELVRCIRFDNKRIVSGAYDGKI 26
DB 422 EGHEELVRCIRFDEKRIIVSGAYDGKI 447

RESULT 5
ID POBRS_SCHPO STANDARD; PRT; 506 AA.
AC 009855; Q9P7V1.
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof11.
GN POF11 OR SPAC29E6.01 OR SPAC30.05.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA Hartung C.L., Toda T.;
RT "Systematic genome-wide analysis of F-box protein-encoding genes in
RL fission yeast."
RN Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Moorey P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Glynnoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Mambuli R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Motlier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

```

```

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Harell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC - SIMILARITY: CONTAINS 7 WD_REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB061694; BAB5543.1; -
DR EMBL: AL136538; CAB66464.1; -
DR EMBL: Z66525; CA91423.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 3.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 8.
DR PROSITE: PS00181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT DOMAIN 70 116 F-BOX.
FT REPEAT 219 256 WD 1.
FT REPEAT 259 298 WD 2.
FT REPEAT 301 338 WD 3.
FT REPEAT 345 386 WD 4.
FT REPEAT 388 426 WD 5.
FT REPEAT 427 464 WD 6.
FT REPEAT 468 505 WD 7.
SO SEQUENCE 506 AA; 58257 MW; CEF34D4EFBC2E10 CRC64;

Query Match
Best Local Similarity 66.7%; Score 92; DB 1; Length 506;
Matches 14; Conservative 53.8%; Pred. No. 1,7e-06; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGHEELVRCIRFDNKRIVSGAYDGKI 26
DB 426 QGHEDLIRTVRFNDEKIVSGYDGTV 451

RESULT 6
ID SE10_CAEEL STANDARD; PRT; 579 AA.
AC Q93794;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sel-10 protein.
GN SEL-10 OR F55B12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sims M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.

```

```

CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U21220; AAC15905.1; -.
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00400; WD40; 7.
CC Pfam: PF00646; F-box; 1.
CC PRINTS: PR00320; GPROTEINBRPT.
CC ProDom: PD000018; WD40; 4.
CC SMART: SM00256; FBOX; 1.
CC SMART: SM00320; WD40; 7.
CC PROSITE: PS00181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS.1; 4.
CC PROSITE: PS50082; WD_REPEATS.2; 7.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC Transcription regulation; Repeat; WD repeat.
CC KW DOMAIN 178 224 F-box.
CC FT REPEAT 347 375 WD 1.
CC FT REPEAT 387 415 WD 2.
CC FT REPEAT 427 455 WD 3.
CC FT REPEAT 466 496 WD 4.
CC FT REPEAT 508 543 WD 5.
CC FT REPEAT 553 595 WD 6.
CC FT REPEAT 607 635 WD 7.
CC FT REPEAT 647 675 WD 8.
CC SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;
CC -----
OY 2 GHPELVRCIRFDKNRIVSGAYDGI 26
Db 387 GHESGIRCIQFDFTKLISGMDRTI 411
| | | : | | | | : | | | | : | |
RESULT 8
APAF_BRARE STANDARD; PRT: 1261 AA.
AC G919H8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Apoptotic protease activating factor 1 (Apaf-1).
GN APAF1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
CX NCBI_TaxID=7955;
CX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT Zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
CC -!- FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent
CC autocatalytic activation of pro-caspase 9 (Apaf-3), leading to the
CC activation of caspase-3 and apoptosis. This activation requires
CC ATP (By similarity).
CC -!- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and
CC dATP (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: CONTAINS 1 CARD DOMAIN.
CC -!- SIMILARITY: CONTAINS 13 WD_REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: CONTAINS 1 NB-ARC DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright It is produced through a collaboration

```



RA Mooney P., Moule S., Murgall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares K.,  
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymopiez B.,  
 RA Woodjens I., Vanstreels E., Rieger M., Scheer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritsch C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Coffau A., Cadiou E., Dreano S., Gloux S., Lelaive V., Motlier S.,  
 RA Calbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
 RA Dugas R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Sprakowski G.V., Usery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AB032410; BAA84528.1; -  
 DR EMBL: Z94864; CAB08168.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 5.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ub1 conjugation; Repeat; WD repeat.  
 KW DOMAIN 107 153  
 FT REPEAT 271 299 WD 1.  
 FT REPEAT 311 339 WD 2.  
 FT REPEAT 350 379 WD 3.  
 FT REPEAT 390 420 WD 4.  
 FT REPEAT 432 460 WD 5.  
 FT REPEAT 472 500 WD 6.  
 FT REPEAT 510 538 WD 7.  
 SQ SEQUENCE 605 AA: 67110 MW: 7118C9379EC5C1F0 CRC64;

Query Match 45.7%; Score 63; DB 1; Length 605;  
 Best Local Similarity 42.3%; Pred. No. 0.055;  
 Matches 11; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 EGHLELYRCIRFD-NKRIYSGAYDGI 26  
 DB 310 EGHSSGVCTLPDCKLTISGSMKRTI 335

RESULT 11  
 WD66.PHYPO STANDARD: PRT: 601 AA.  
 AC P90587;  
 DT 15-JUL-1998 (rel. 36, Created)

DT 15-JUL-1998 (rel. 36, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE 66 kDa stress protein (p66).  
 OS Physarum polycephalum (Slime mold).  
 OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;  
 OC Physarum.  
 OX NCBI\_TaxID=5791;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98351989; PubMed=9685722;  
 RA Matsumoto S., Ogawa M., Kasakura T., Shimada Y., Mitsui M., Maruya M.,  
 RA Isobata M., Yahara I., Murakami-Murofushi K.,  
 RA "A novel 66-kDa stress protein, p66, associated with the process of  
 RT cyst formation of Physarum polycephalum is a Physarum homologue of a  
 RT yeast actin-interacting protein, Aip1."  
 RL J. Biochem. 124:326-331(1998).  
 CC -1- FUNCTION: ASSOCIATED WITH THE PROCESS OF CYST FORMATION.  
 CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE AIP1 FAMILY OF WD-REPEAT PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U86011; AAC26321.1; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 10.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR SMART: SM00320; WD40; 10.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; FALSE\_NEG.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT REPEAT 56 95 WD 1.  
 FT REPEAT 100 143 WD 2.  
 FT REPEAT 145 184 WD 3.  
 FT REPEAT 187 226 WD 4.  
 FT REPEAT 233 272 WD 5.  
 FT REPEAT 318 357 WD 6.  
 FT REPEAT 357 395 WD 7.  
 FT REPEAT 435 478 WD 8.  
 FT REPEAT 483 522 WD 9.  
 FT REPEAT 526 565 WD 9.  
 FT REPEAT 569 600 WD 10.  
 SQ SEQUENCE 601 AA: 64321 MW: F691217D838F747A CRC64;

Query Match 44.9%; Score 62; DB 1; Length 601;  
 Best Local Similarity 46.4%; Pred. No. 0.078;  
 Matches 13; Conservative 6; Mismatches 7; Indels 2; Gaps 1;

QY 1 EGHLELYRCIRFD-NKRIYSGAYDGI 26  
 DB 317 KGHNKLVTSIAFDTSKALYSIDYGV 344

RESULT 12  
 MT30.YEAST STANDARD: PRT: 640 AA.  
 AC P39014;  
 DT 01-FEB-1995 (rel. 31, Created)  
 DT 01-FEB-1995 (rel. 31, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE MET30 protein.  
 DE MET30 OR YIL046W.  
 GN Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-X2180-1A;  
 MEDLINE-96069360; PubMed-8524217;  
 RA Thomas D., Kuras L., Barbery R., Cherest H., Blaiseau P.L.,  
 RA Surdin-Kerjan Y.,  
 RT "Wet30p, a yeast transcriptional inhibitor that responds to S-  
 RT adenosylmethionine, is an essential protein with WD40 repeats.",  
 RL Mol. Cell. Biol. 15:6526-6534(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288c / AB972;  
 RA Bartell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
 RA Chercher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis E., Iye G., Moule S., Moule T., Odell C., Pearson D.,  
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.,  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS  
 CC GENES EXPRESSION.  
 CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCOMB/SCON-2 FAMILY OF WD-REPEAT  
 CC PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 246861; CAA86905.1; -;  
 DR EMBL: 126505; AAA96717.1; -;  
 DR SGD: S0001308; MEM30.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Transcription regulation; Methionine biosynthesis;  
 KW Cysteine biosynthesis; Repeat; WD repeat.  
 FT DOMAIN 181 227  
 FT REPEAT 300 328 WD 1.  
 FT REPEAT 340 368 WD 2.  
 FT REPEAT 380 408 WD 3.  
 FT REPEAT 419 449 WD 4.  
 FT REPEAT 461 499 WD 5.  
 FT REPEAT 509 538 WD 6.  
 FT REPEAT 550 578 WD 7.  
 FT REPEAT 607 635 WD 8 (POTENTIAL).  
 FT CONFLICT 61 61 M -> I (IN REF. 1).  
 SO SEQUENCE 640 AA; 72835 MW; 5135D4BCA2E1B97 CRC64;  
 QY Query Match 44.9%; Score 62; DB 1; Length 640;  
 Best Local Similarity 56.0%; Pred. No. 0.083;  
 Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

RESULT 13  
 ATPL\_YEAST

ID ATPL\_YEAST STANDARD; PRT; 615 AA.  
 AC P46860;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Actin interacting protein 1.  
 GN ATPL OR YMR092C OR YW582.17C.  
 OS Saccharomyces cerevisiae (baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxId=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Amberg D.C., Botstein D.,  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288c / AB972;  
 RA Gentles S., Bowman S., Bartell B.G., Rajandream M.A.,  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP INTERACTION WITH COFILIN, AND FUNCTION.  
 RX PubMed-10366597;  
 RA Rodal A.A., Tetreault J.W., Lappalainen P., Drubin D.G., Amberg D.C.,  
 RT J. Cell Biol. 145:1251-1264(1999).  
 RL  
 CC -1- FUNCTION: Involved in the depolymerization of actin filaments.  
 CC Enhances the filament disassembly activity of cofilin and  
 CC restricts cofilin localization to cortical actin patches.  
 CC -1- SUBUNIT: Interacts with actin and cofilin in a ternary complex.  
 CC -1- SIMILARITY: CONTAINS 11 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE ATPL FAMILY OF WD-REPEAT PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U35666; AAA79141.1; -;  
 DR EMBL: 249259; CAA89239.1; -;  
 DR SGD: S0004698; ATPL.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; FALSE\_NEG.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 2.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 2.  
 KW Actin-binding; Cytoskeleton; Repeat; WD repeat.  
 FT REPEAT 62 109 WD 1.  
 FT REPEAT 114 157 WD 2.  
 FT REPEAT 158 201 WD 3.  
 FT REPEAT 204 244 WD 4.  
 FT REPEAT 251 289 WD 5.  
 FT REPEAT 337 367 WD 6.  
 FT REPEAT 370 406 WD 7.  
 FT REPEAT 442 481 WD 8.  
 FT REPEAT 486 525 WD 9.  
 FT REPEAT 529 578 WD 10.  
 FT REPEAT 583 613 WD 11.  
 FT DOMAIN 363 366 POLY-SER.  
 SO SEQUENCE 615 AA; 67325 MW; B8340BF68DD08257 CRC64;  
 QY Query Match 44.2%; Score 61; DB 1; Length 615;  
 Best Local Similarity 39.4%; Pred. No. 0.11;  
 Matches 13; Conservative 7; Mismatches 5; Indels 8; Gaps 1;

QY 2 GHDELVRCIRFDNKR-----VSGAYDGI 26  
 DB 327 GHDELVKTISGHNKGITATLVNPLISGSDGRI 359



RESULT 14  
SC02\_NEUCR STANDARD: PRT: 650 AA.  
ID SC02\_NEUCR  
AC 001277.  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sulfur controller-2 (SCON2).  
GN SCON-2.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A;  
RX MEDLINE=95241499; PubMed=7724564;  
RA Kumar A., Paletta J.V.;  
RT "The sulfur controller-2 negative regulatory gene of Neurospora  
crassa encodes a protein with beta-transducin repeats.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).  
CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.  
CC -1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.  
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT  
PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U17251; AAA68968.1;  
DR InterPro: IPR001810; F-box.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF004400; WD40; 7.  
DR Pfam: PF00646; F-box; 1.  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR PRODOM: PD000018; WD40; 4.  
DR SMART: SM00256; FBOX; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS00181; FBOX; 1.  
DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
DR PROSITE: PS00082; WD\_REPEATS\_2; 6.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Transcription regulation; Repeat; WD repeat.  
KW DOMAIN  
FT 124 170 F-BOX.  
FT REPEAT 292 320 WD 1.  
FT REPEAT 332 360 WD 2.  
FT REPEAT 372 400 WD 3.  
FT REPEAT 411 441 WD 4.  
FT REPEAT 453 488 WD 5.  
FT REPEAT 528 564 WD 6.  
FT REPEAT 576 604 WD 7.  
FT REPEAT 616 644 WD 8.  
SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 42.8%; Score 59; DB 1; Length 650;  
Best Local Similarity 38.5%; Pred. No. 0.24;  
Matches 10; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 1 EGHEELVRCIRFDNKRIVGAYDGKI 26  
DB 291 KGHNGVTCLODDNIIATGSDYDTI 316

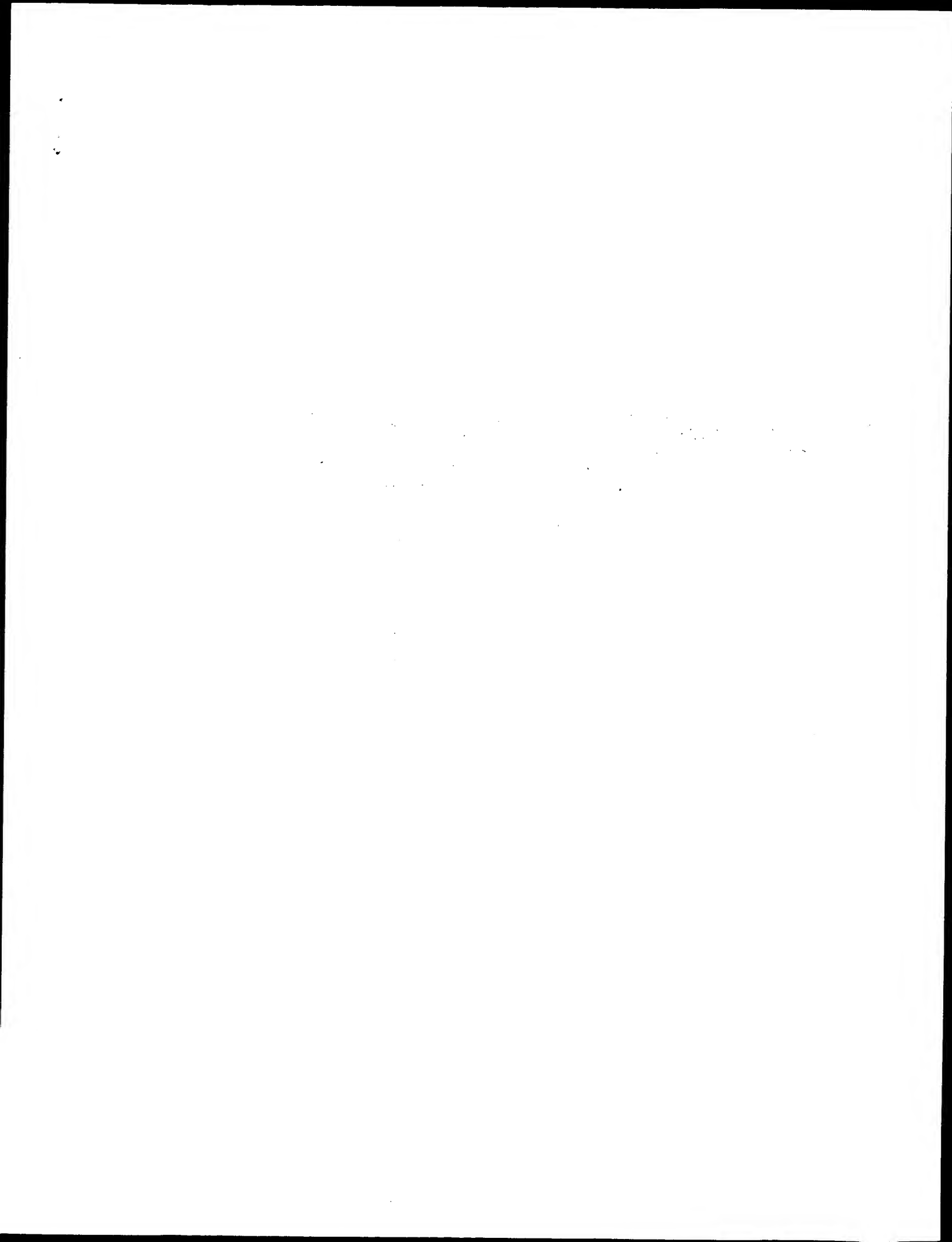
RESULT 15

YK16\_YEAST  
ID YK16\_YEAST STANDARD: PRT: 659 AA.  
AC P36130;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 74.7 kDa Trp-Asp repeats containing protein in DAL80-GAP1  
intergenic region.  
GN YK036C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Urrestazu L.A., Jauniaux J.-C.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z28261; CAA82110.1;  
DR PIR: S38108; S38108.  
DR TRASFAC: T04326;  
DR SGI: S0001744; YK036C.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF004400; WD40; 6.  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR PRODOM: PD000018; WD40; 1.  
DR SMART: SM00320; WD40; 5.  
DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 5.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Hypothetical protein; Repeat; WD repeat.  
FT REPEAT 333 364 WD 1.  
FT REPEAT 376 407 WD 2.  
FT REPEAT 438 466 WD 3.  
FT REPEAT 503 531 WD 4.  
FT REPEAT 543 571 WD 5.  
FT REPEAT 582 610 WD 6.  
SQ SEQUENCE 659 AA; 74709 MW; 139593554C56DD37 CRC64;

Query Match 42.8%; Score 59; DB 1; Length 659;  
Best Local Similarity 30.8%; Pred. No. 0.25;  
Matches 8; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

OY 1 EGHEELVRCIRFDNKRIVGAYDGKI 26  
DB 542 EGHITDGTSLKFDSEKLVGMSDNVS 567

Search completed: April 10, 2003, 13:19:38  
Job time : 2.86845 secs



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 ; Search time 6.14373 Seconds  
(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_467\_492  
Perfect score: 138  
Sequence: 1 EGHELVRCIRFDNKRIVSGAYDGKI 26

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.podent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	138	100.0	252	11	0922C7 mus musculu
2	138	100.0	563	11	0923H0 mus musculu
3	138	100.0	569	11	092159 mus musculu
4	138	100.0	569	11	090015 mus musculu
5	133	96.4	430	5	09BJ54 mus musculu
6	133	96.4	569	11	09RIG7 mus musculu
7	132	95.7	510	5	044382 drosophila
8	132	95.7	510	5	09VDE3 drosophila
9	82	59.4	553	4	09NUX6 homo sapien
10	82	59.4	561	4	096R12 homo sapien
11	82	59.4	589	4	096LE0 homo sapien
12	82	59.4	627	4	096A16 homo sapien
13	82	59.4	629	11	08VAP4 mus musculu
14	82	59.4	629	11	08VAP4 mus musculu
15	82	59.4	707	4	0969H0 homo sapien
16	78	56.5	1326	5	09VZF4 drosophila

17	74	53.6	585	5	095ZT0	095ZT0 caenorhabd
18	74	53.6	587	5	044083	044083 caenorhabd
19	68	49.3	1101	5	062471	062471 caenorhabd
20	65	47.1	329	5	08SSP4	08SSP4 dictyostell
21	64	46.4	196	4	095611	095611 homo sapien
22	64	46.4	252	4	043848	043848 homo sapien
23	64	46.4	252	4	095170	095170 homo sapien
24	64	46.4	999	4	09H0F0	09H0F0 homo sapien
25	64	46.4	1061	4	08TC00	08TC00 homo sapien
26	63	45.7	423	5	08SSJ1	08SSJ1 encephalito
27	61	44.2	178	4	09H7B8	09H7B8 homo sapien
28	61	44.2	1883	4	09H2Y7	09H2Y7 homo sapien
29	59	42.8	176	11	09ESU3	09ESU3 mus musculu
30	59	42.8	339	11	088465	088465 mus musculu
31	59	42.8	562	10	09SC73	09SC73 oryza saliv
32	59	42.8	1888	11	088466	088466 mus musculu
33	58.5	42.4	589	5	096661	096661 trypanosoma
34	58	42.0	297	5	09VE10	09VE10 drosophila
35	58	42.0	357	16	08YUJ4	08YUJ4 anabaena sp
36	58	42.0	514	10	094IP9	094IP9 lycopersico
37	58	42.0	680	5	09V5W2	09V5W2 drosophila
38	57	41.7	386	5	09XWU3	09XWU3 caenorhabd
39	57	41.3	440	10	004634	004634 arabidopsis
40	57	41.3	511	10	08W117	08W117 arabidopsis
41	57	41.3	522	10	09C9T9	09C9T9 thermocoga
42	57	41.3	580	16	09X2G1	09X2G1 thermocoga
43	56.5	40.9	494	3	094365	094365 schizosacch
44	56.5	40.9	586	3	060136	060136 schizosacch
45	56.5	40.9	774	5	09V767	09V767 drosophila

## ALIGNMENTS

## RESULT 1

0922C7 PRELIMINARY: PRT: 252 AA.

AC 0922C7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Similar to f-box and WD-40 domain protein 1B (Fragment).  
GN FBXW1B.  
OS Mus musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (May-2001) to the EMBL/Genbank/DDAJ databases.  
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
DR EMBL: BC008552; AAH08552.1; -  
DR MGI: 2144023; Fbxw1b.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 6.  
DR PRINTS: PR00320; GPROTEINBPT.  
DR PRODOM: PD000018; WD40; 3.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
DR PROSITE: PS00082; WD\_REPEATS\_2; 6.  
DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
FT NON\_TER  
SQ  
SD  
STOUCENCE 252 AA; F71737C8BD7A9F75F CRC64;

Query Match 100.0%; Score 138; DB 11; Length 252;  
Best Local Similarity 100.0%; Pred. No. 2.1e-13;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHELVRCIRFDNKRIVSGAYDGKI 26  
DB 150 EGHELVRCIRFDNKRIVSGAYDGKI 175

```

RESULT 2
ID 0923H0 PRELIMINARY; PRT; 563 AA.
AC 0923H0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE F-box/WD40 repeat-containing protein HOS.
GN FBXW1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RA Bhalla N., Hetter J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;
RT "Mouse homolog of HOS (mhOS) is overexpressed in skin tumors and
RT implicated in constitutive activation of NF-kappaB.";
CC Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AY038079; AKK72095.1; -.
DR MGD: MGI:2144023; FBxw1b.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR ProDom: PD000018; WD40; 7.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN. 5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 563 AA; 64741 MW; 9AB562F3FE5E3496 CRC64;

Query Match 100.0%; Score 138; DB 11; Length 563;
Best Local Similarity 100.0%; Pred. No. 4.9e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHLELVRCIRFDNKRIVSGAYDGKI 26
DB 461 EGHLELVRCIRFDNKRIVSGAYDGKI 486

RESULT 3
ID 092159 PRELIMINARY; PRT; 569 AA.
AC 092159;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Beta-transducin repeat containing protein.
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99145465; PubMed=9990853;
RA Spencer E., Jiang J., Chen Z.J.;
RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT Slmb/beta-Trcp.";
RL Genes Dev. 13:284-294(1999).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF112979; AAD04181.1; -.
DR MGD: MGI:1338871; Btrc.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR ProDom: PD000018; WD40; 7.
DR PRINTS: PR00320; GPROTEINBRPT.

```

```

DR ProDom: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN. 6.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;

Query Match 100.0%; Score 138; DB 11; Length 569;
Best Local Similarity 100.0%; Pred. No. 4.9e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHLELVRCIRFDNKRIVSGAYDGKI 26
DB 467 EGHLELVRCIRFDNKRIVSGAYDGKI 492

RESULT 4
ID 090U15 PRELIMINARY; PRT; 569 AA.
AC 090U15;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)
DE (F-box-WD40 repeat protein 1).
GN BTRC OR FBXW1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9919275; PubMed=10097128;
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hattori K., Higashi H., Nakano H., Okumura K., Ono K., Good R.A.,
RA Nakayama K.-I.;
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT ubiquitin ligase Skp1/Cul1/F-box protein FWD1.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99075339; PubMed=9859996;
RA Varon A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase.";
RL Nature 396:590-594(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SURAIN=129/SV;
RA MEDLINE=21601157; PubMed=11735228;
RA Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,
RA Nakayama K.-I.;
RT "Characterization of a Mouse Gene (FBXW6) That Encodes a Homologue of
RT Caenorhabditis elegans SEL-10.";
RL Genomics 78:214-222(2001).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF081887; AAD17755.1; -.
DR EMBL: AF099932; AAD08701.1; -.
DR EMBL: BC003989; AAO03989.1; -.
DR EMBL: AF391190; AAL40929.1; -.
DR EMBL: AF391178; AAL40929.1; JOINED.
DR EMBL: AF391179; AAL40929.1; JOINED.
DR EMBL: AF391180; AAL40929.1; JOINED.
DR EMBL: AF391181; AAL40929.1; JOINED.
DR EMBL: AF391182; AAL40929.1; JOINED.

```

DR EMBL: AF391183; AAL40929.1; JOINED.  
 DR EMBL: AF391184; AAL40929.1; JOINED.  
 DR EMBL: AF391185; AAL40929.1; JOINED.  
 DR EMBL: AF391186; AAL40929.1; JOINED.  
 DR EMBL: AF391187; AAL40929.1; JOINED.  
 DR EMBL: AF391188; AAL40929.1; JOINED.  
 DR EMBL: AF391189; AAL40929.1; JOINED.  
 DR MGD: MGI:1338871; Btcr.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Ligase; Repeat; WD repeat.  
 KW SEQUENCE 569 AA; 65105 MW; BCD6544815B2296 CRC64;

Query Match 100.0%; Score 138; DB 11; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-13;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHEELVRCIRFDNKRIVSGAYDGKI 26  
 DB 467 EGHEELVRCIRFDNKRIVSGAYDGKI 492

RESULT 5  
 Q9BJ54 PRELIMINARY: PRT: 430 AA.  
 AC Q9BJ54: 09BJ54: 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)  
 DE Beta-transducin repeat-containing protein (fragment).  
 OS Heterodera glycyines (Soybean cyst nematode).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida;  
 OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera;  
 NCBI\_Taxid=51029;

RA KOVALEVA E.S., YAKOVLEV A.G., MASLER E.P.;  
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF391101; AAK26376.1; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER 1  
 SQ SEQUENCE 430 AA; 47916 MW; 4ABC3FF2DFE3A50B CRC64;

Query Match 96.4%; Score 133; DB 5; Length 430;  
 Best Local Similarity 96.2%; Pred. No. 2.2e-12;  
 Matches 25; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGHEELVRCIRFDNKRIVSGAYDGKI 26  
 DB 291 EGHEELVRCIRFDNKRIVSGAYDGKI 316

RESULT 6

O9RIG7 PRELIMINARY: PRT: 569 AA.  
 AC O9RIG7: 09RIG7: 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)  
 DE Beta-transducin repeat-containing protein.  
 GN BTRC.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Winston J., Ellledge S.J., Harper J.W.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF110396; AAD41025.1; -  
 DR MGD: MGI:1338871; Btcr.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Repeat; WD repeat.  
 KW SEQUENCE 569 AA; 65209 MW; E6DDCAD28D5519D CRC64;

Query Match 96.4%; Score 133; DB 11; Length 569;  
 Best Local Similarity 96.2%; Pred. No. 3e-12;  
 Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGHEELVRCIRFDNKRIVSGAYDGKI 26  
 DB 467 EGHEELVRCIRFDNKRIVSGAYDGKI 492

RESULT 7  
 O44382 PRELIMINARY: PRT: 510 AA.  
 AC O44382: 044382: 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)  
 DE SLIMB.  
 GN SLIMB OR SLIMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:98121115; PubMed:9461217;  
 RA Jiang J., Struhl G.;  
 RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb.";  
 RL Nature 391:493-496(1998).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF032878; AAC38852.1; -  
 DR FlyBase: FBgn0023423; slimb.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.



```

RP RP SEQUENCE FROM N.A., Myers R.L., Shuang R., Brashler J.R., Yan R.,
RA Li J., Pauley A.M., Gunney M.E.;
RA Buhl A.F., Gunney M.E.;
RT "SEL-10 interacts with Presenilin 1, facilitates its ubiquitination,
RT and Alters A-beta production";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AY008274; AAC16640.1; -.
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR ProDom: PD00018; WD40; 7.
DR PROSITE: PS50181; FBOX; 2.
DR PROSITE: PS500678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KM Repeat: WD repeat.
SO SEQUENCE 589 AA; 66120 MW; 2AFB68A36BE8DE CRC64;

Query Match          59.4%; Score 82; DB 4; Length 589;
Best Local Similarity 59.1%; Pred. No. 0.00027;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0.

OY      2 GHLELVRCIRFDNKRIVSGAYD 23
        |||::|||::|||::|||::|||
Db       381 GHVAAVRCVQYDGRRVVSGAYD 402

RESULT 12
ID   O96A16             PRELIMINARY;           PRT:     627 AA.
AC   O96A16;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE   Archipelago beta form (F-box protein FBW7).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxId=9606;
RN   [1]
RP   MEDLINE=21444047; PubMed=11565033;
RA   Moberg K.H., Bell D.W., Maher D.C., Haber D.A., Haritharan I.K.;
RT   "Archipelago regulates Cyclin E levels in Drosophila and is mutated in
RL   human cancer cell lines";
RL   Nature 413:311-316(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20003061; PubMed=10531037;
RA   Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.;
RT   "A family of mammalian F-box proteins.";
RL   Curr. Biol. 9:1180-1182(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Koepf D.M., Winston J.T., Harper W., Ellledge S.J.;
RT   Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC   -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR   EMBL: AF411972; AAL06291.1; -.
DR   EMBL: AY033553; AAK57347.1; -.
DR   InterPro: IPR001810; F-box.
DR   InterPro: IPR001680; WD40.
DR   Pfam: PF00646; F-box; 1.
DR   Pfam: PF00400; WD40; 7.
DR   ProDom: PD000018; WD40; 2.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KM Repeat: WD repeat.
SO SEQUENCE 627 AA; 70324 MW; 3DA107C05318BED CRC64;
```

Query Match 59.4%; Score 82; DB 4; Length 627;  
Best Local Similarity 59.1%; Pred. No. 0.00029;  
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23  
DB 419 GHVAARVCYOYDGRVVSAYD 440

## RESULT 13

Q8VHP4 PRELIMINARY; PRT; 629 AA.

AC Q8VHP4; 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE F-box protein.

GN PF00646; F-box; 1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Ilyin G.P.; Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF427101; AAL50052.1; -

DR InterPro; IPR001810; F-box.

DR InterPro; IPR001680; WD40.

DR Pfam; PF00646; F-box; 1.

DR PRINTS; PR00320; GPROTEINRPT.

DR PRODOM; PD000018; WD40; 2.

DR SMART; SM00256; FBOX; 1.

DR PROSITE; PS50181; FBOX; 1.

DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_5.

DR PROSITE; PS50082; WD\_REPEATS\_2; 7.

DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.

DR SQUONCE 629 AA; 70562 MW; BE916405A3490A3E CRC64;

Query Match 59.4%; Score 82; DB 11; Length 629;  
Best Local Similarity 59.1%; Pred. No. 0.00029;

Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23  
DB 421 GHVAARVCYOYDGRVVSAYD 442

## RESULT 14

Q8VHP4 PRELIMINARY; PRT; 629 AA.

AC Q8VHP4; 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE F-box WD40 repeat protein 6.

GN FBXW6.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RA MEDLINE=129/SV.

RA MEDLINE=21601157; PubMed=11735228;

RA Maryama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,

RA Nakayama K.-I.; Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of

RT Caenorhabditis elegans SEL-10.";

RL Genomics 78:214-222(2001).

DR EMBL; AF391202; AAL40930.1; -

DR EMBL; AF391193; AAL40930.1; JOINED.  
DR EMBL; AF391194; AAL40930.1; JOINED.  
DR EMBL; AF391195; AAL40930.1; JOINED.  
DR EMBL; AF391196; AAL40930.1; JOINED.  
DR EMBL; AF391197; AAL40930.1; JOINED.  
DR EMBL; AF391198; AAL40930.1; JOINED.  
DR EMBL; AF391199; AAL40930.1; JOINED.  
DR EMBL; AF391200; AAL40930.1; JOINED.  
DR EMBL; AF391201; AAL40930.1; JOINED.  
DR EMBL; AF391202; AAL40930.1; JOINED.  
DR InterPro; IPR001810; F-box.  
DR InterPro; IPR001680; WD40.  
DR Pfam; PF00646; F-box; 1.  
DR PRINTS; PR00320; GPROTEINRPT.  
DR PRODOM; PD000018; WD40; 2.  
DR SMART; SM00256; FBOX; 1.  
DR PROSITE; PS50181; FBOX; 1.  
DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
DR SQUONCE 629 AA; 70547 MW; 80483A7C2ED61FE7 CRC64;

Query Match 59.4%; Score 82; DB 11; Length 629;  
Best Local Similarity 59.1%; Pred. No. 0.00029;  
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23  
DB 421 GHVAARVCYOYDGRVVSAYD 442

## RESULT 15

Q969H0 PRELIMINARY; PRT; 707 AA.

AC Q969H0; 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE F-box protein CDC4 (Archipelago alpha form).

GN CDC4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RA MEDLINE=21449048; PubMed=11565034;

RA Strohmaier H., Spruck C.H., Kaiser P., Won K.A., Sangfelt O.,

RA Reed S.L.; "Human F-box protein hcdc4 targets cyclin E for proteolysis and is

RT mutated in a breast cancer cell line.";

RL Nature 413:316-322(2001).

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=21449047; PubMed=11565033;

RA Moberg K.H., Bell D.W., Wahner D.C., Haber D.A., Hariharan I.K.;

RT "Archipelago regulates cyclin E levels in Drosophila and is mutated in

RL human cancer cell lines.";

CC -1- SIMILARITY: CONTAINS 7 WD\_REPEATS (TRP-ASP DOMAINS).

DR EMBL; AY049984; AAL07271.1; -

DR EMBL; AF411971; AAL06290.1; -

DR InterPro; IPR001810; F-box.

DR InterPro; IPR001680; WD40.

DR Pfam; PF00646; F-box; 1.

DR PRODOM; PD000018; WD40; 2.

DR PROSITE; PS50181; FBOX; 1.

DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_5.

DR PROSITE; PS50082; WD\_REPEATS\_2; 7.

DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.

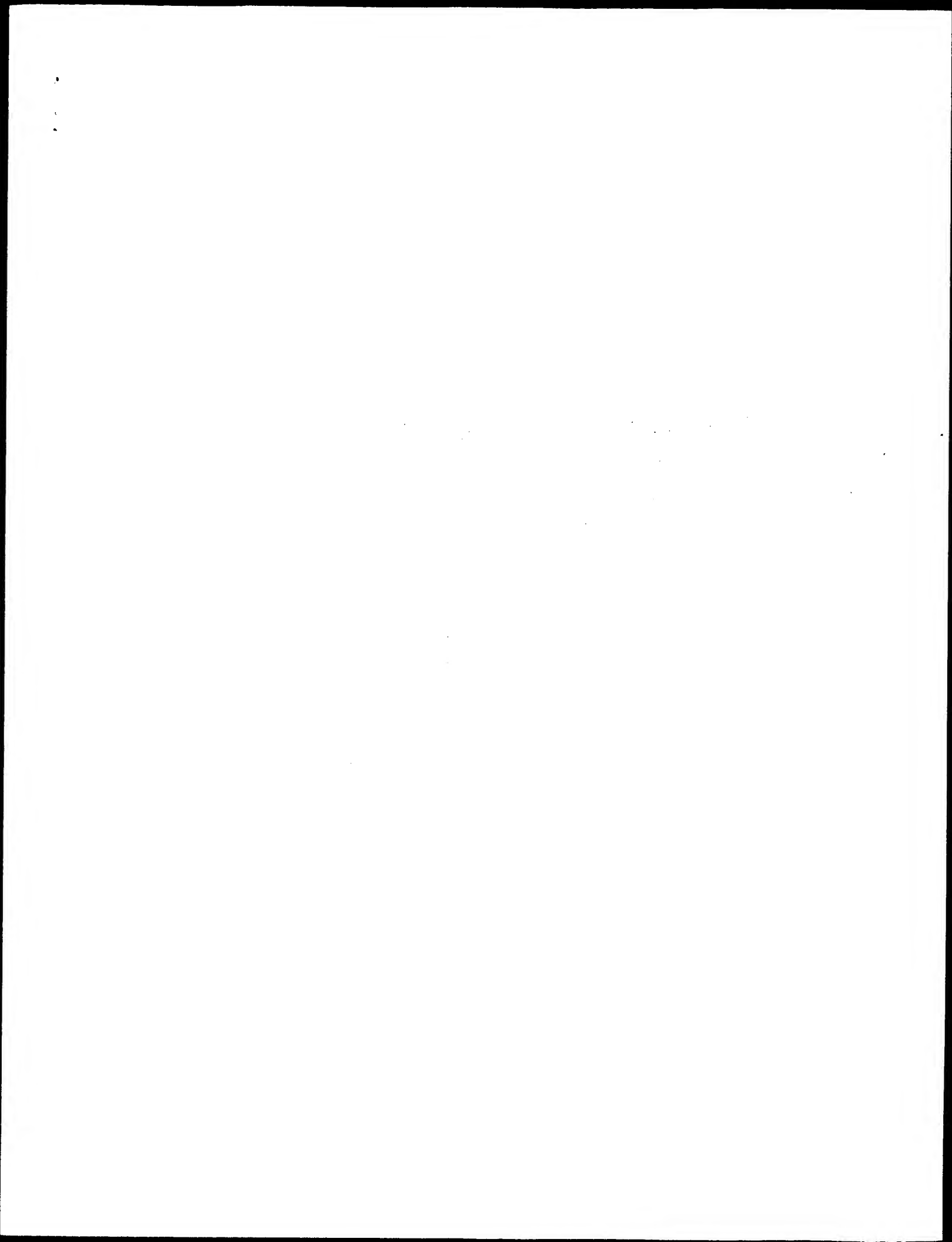


KW Repeat: WD repeat.  
SQ SEQUENCE 707 AA; 79663 MW; EAA357F76DFD8203 CRC64;

Query Match 59.4%; Score 82; DB 4; Length 707;  
Best Local Similarity 59.1%; Pred. No. 0.00033;  
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 GHEELVRCIRFDNKRIVSGAYD 23  
|||:::|:|::|  
Db 499 GHVAAVRCVQYDGRVVGAYD 520

Search completed: April 11, 2003, 11:51:54  
Job time : 8.14373 secs



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: April 11, 2003, 11:20:44 ; Search time 8.93666 Seconds  
(without alignments)  
432.406 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_516\_544  
Perfect score: 152  
Sequence: 1 VEHSGRVFRLQFDEQIVSSHDDTLIN 29

Scoring table: BLOSUM62  
Gapex 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	100.0	517	16 AAR5852	WD-40 domain-contg
2	152	100.0	569	20 AAR54054	A human beta-trans
3	152	100.0	569	21 AAB12812	Mouse ubiquitin 11
4	152	100.0	569	21 AAB12813	Human beta-transdu
5	152	100.0	569	21 AAY96697	Human beta-TRCP.
6	152	100.0	569	21 AAY83041	F-box protein FBP
7	152	100.0	569	21 AAY83250	F-box protein hbet
8	152	100.0	569	21 AAY83254	F-box protein FWD1
9	152	100.0	569	21 AAY44249	Human cell signall
10	152	100.0	569	22 AAB48298	Human ZP11 protein

11	152	100.0	579	22 AAM78583	Human protein SEQ
12	152	100.0	590	22 AAM00847	Human bone marrow
13	152	100.0	605	22 AAM78582	Human protein SEQ
14	152	100.0	608	22 AAM00960	Human bone marrow
15	152	100.0	632	22 AAM78584	Human protein SEQ
16	152	100.0	654	22 AAM79566	Human protein SEQ
17	152	100.0	654	22 AAM79567	Human protein SEQ
18	152	100.0	654	22 AAM79568	Human protein SEQ
19	151	99.3	542	21 AAY96696	Human E3 ubiquitin
20	151	99.3	542	21 AAM79127	Human protein SEQ
21	151	99.3	542	22 AAM40208	Human polypeptide
22	151	99.3	550	22 AAM41994	Human polypeptide
23	149	98.0	510	16 AAB59857	Drosophila melanog
24	148	97.4	29	16 AAR84941	Peptide rvi1 from
25	73	48.0	540	20 AAY22465	Human hippocampal
26	73	48.0	540	22 AAY22468	Human mammary sel-
27	73	48.0	540	22 AAB59197	Human hippocampal
28	73	48.0	540	22 AAB59200	Human mammary sel-
29	73	48.0	545	20 AAY22464	Human hippocampal
30	73	48.0	545	22 AAB59196	Human hippocampal
31	73	48.0	553	22 AAY22463	Human hippocampal
32	73	48.0	553	22 AAB93475	Human protein sequ
33	73	48.0	553	22 AAB59195	Human hippocampal
34	73	48.0	559	20 AAY22467	Human mammary sel-
35	73	48.0	559	22 AAB59199	Human mammary sel-
36	73	48.0	589	20 AAY22466	Human mammary sel-
37	73	48.0	589	21 AAB01204	Human GTPase assoc
38	73	48.0	589	22 AAB59198	Human mammary sel-
39	73	48.0	592	20 AAY22462	Human hippocampal
40	73	48.0	592	22 AAB59194	Human hippocampal
41	73	48.0	626	20 AAY22469	Human fmyc-N-sel-1
42	73	48.0	626	22 AAB59201	Protein encoded by
43	73	48.0	627	20 AAY22461	Human hippocampal
44	73	48.0	637	22 AAB59193	Human hippocampal
45	73	48.0	666	20 AAY22471	Human C-term mychi

## ALIGNMENTS

RESULT 1	
AA85852	
ID AAR85852 standard; peptide: 517 AA.	
XX	
XX AAR85852:	
XX	
XX 13-SEP-1996 (first entry)	
XX	
DE WD-40 domain-contg. beta-TRCP protein.	
XX	
KW WD40 repeat region; beta-transducin; protein-protein interaction; drug;	
KW intracellular signalling; protein kinase C; homology; motif; modulator;	
KW receptors of activated protein kinase; enzyme activity; isozyme; human.	
XX	
OS Synthetic.	
XX	
PN WO9521252-A2.	
XX	
PD 10-AUG-1995.	
XX	
PF 31-JAN-1995: 95WO-US01210.	
XX	
PR 01-FEB-1994: 94US-0190802.	
XX	
PA (STRD ) UNIV LELAND STANFORD JUNIOR.	
XX	
PI Mochly-Rosen D, Ron D;	
XX	
XX WPL: 1995-283772/37.	
XX	
DR New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the	
PT activity of a protein, eg. protein kinase C, which interacts with a	
PT protein contg. a WD-40 region.	

PR	30-JAN-1998;	98FR-0001100.	
XX			
PA	(INRM ) INSERM INST NAT SANTE & RECH MEDICALE.		
PA	(INSP ) INST PASTEUR.		
XX			
PI	Arenzana Selsdedos F, Benarous R, Concordet J, Durand H;		
PI	Kroll M, Margottin F;		
XX			
DR	WP1: 1999-469329/39.		
DR	N-PSDB: AAX86501.		
XX			
PT	New human beta-transducin repeat containing protein and its		
PT	fragments useful as, or to screen for, antiviral, antitumour,		
XX	anti-inflammatory and anti-Alzheimer's agents		
PS	Claim 1; Page 60-61; 71pp; French.		
XX			
CC	The present sequence represents a human beta-transducin repeat containing		
CC	protein (beta-Trcp). The protein directs proteins to the proteosome		
CC	degradation pathways. The protein is able to interact with the Vpu		
CC	protein of human immune deficiency virus-1 (HIV-1), cellular proteins		
CC	Ikkappa or beta-catenin (bc) and/or protein Skp1. The protein controls		
CC	ubiquitylation of phosphorylated proteins and thus their targeting to		
CC	proteosomes for degradation. Depending on whether the process is		
CC	inhibited or promoted, the result may be delayed breakdown of Cdk4 (in		
CC	cases of HIV-1 infection); increased activity of Ikb (and thus reduced		
CC	activity of NFkappaB) and increased degradation of mutant bc in tumour		
CC	cells, or increased bc survival (and reduced apoptosis) in Alzheimer's		
CC	patients. The beta-Trcp protein, and its active peptide fragments, or its		
CC	nucleic acid, are used to screen for anti HIV-1 agents (antivirals),		
CC	antitumour agents that disrupt cell cycle regulation or protein		
CC	degradation in human tumour cells; and anti-inflammatory agents that		
CC	disrupt activation by NFkappaB. Fragments of the protein are also		
CC	useful for treating osteo-articular inflammation or acute inflammation		
CC	associated with release of tumour necrosis factor.		
XX			
SQ	Sequence 569 AA:		
	Query Match 100.0%; Score 152; DB 20; Length 569;		
	Best Local Similarity 100.0%; Pred. No. 6,7e-15;		
	Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 VEHSGRVRFLQEFDEPQIVSSSHDDITLIM 29		
DB	516 VEHSGRVRFLQEFDEPQIVSSSHDDITLIM 544		
	RESULT 3		
	AAB12812		
	ID AAB12812 standard; protein; 569 AA.		
XX			
AC	AAB12812;		
XX			
XX	27-NOV-2000 (first entry)		
DE			
DE	Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.		
XX			
KM	ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkkappaB;		
KM	beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;		
KM	gene therapy; colon cancer; beta-transducin repeat containing protein;		
XX	beta-Trcp.		
OS			
OS	Mus musculus.		
XX			
PN	JP2000166542-A.		
XX			
PD	20-JUN-2000.		
XX			
PF	02-DEC-1998; 98JP-0343437.		
XX			
PR	02-DEC-1998; 98JP-0343437.		
XX			
XX			
XX	(KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.		

XX WPI: 2000-485550/43.  
 DR N-PSDB; AAA73131.  
 XX F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of IkappaB or beta-catenin  
 XX  
 PS Claim 2; Page 9-10; 19pp; Japanese.  
 XX  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 XX  
 SQ Sequence 569 AA;  
 XX  
 Query Match 100.0%; Score 152; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 VEHSGRVRLQDFEFOIVSSSHDPTILIW 29  
 DB 516 VEHSGRVRLQDFEFOIVSSSHDPTILIW 544  
 XX  
 RESULT 4  
 AAB12813  
 ID AAB12813 standard; protein; 569 AA.  
 XX  
 AC AAB12813;  
 XX  
 DT 27-NOV-2000 (first entry)  
 XX  
 DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.  
 XX  
 KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;  
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KW beta-TrCP.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP200016542-A.  
 XX  
 PD 20-JUN-2000.  
 XX  
 PF 02-DEC-1998; 98JP-0343437.  
 XX  
 PR 02-DEC-1998; 98JP-0343437.  
 XX  
 PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 XX  
 DR WPI: 2000-485550/43.  
 DR N-PSDB; AAA73132.  
 XX  
 PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of IkappaB or beta-catenin  
 XX  
 PS Claim 3; Page 10-12; 19pp; Japanese.  
 XX  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of IkappaB or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TrCP)). The F-box protein can be used for

CC the gene therapy of colon cancer by being recombined to a virus vector.  
 XX  
 SQ Sequence 569 AA;  
 XX  
 Query Match 100.0%; Score 152; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 VEHSGRVRLQDFEFOIVSSSHDPTILIW 29  
 DB 516 VEHSGRVRLQDFEFOIVSSSHDPTILIW 544  
 XX  
 RESULT 5  
 AAY96697  
 ID AAY96697 standard; protein; 569 AA.  
 XX  
 AC AAY96697;  
 XX  
 DT 26-SEP-2000 (first entry)  
 XX  
 DE Human beta-TrCP.  
 XX  
 KW E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KW anti-inflammatory; immunosuppressive; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200034447-A2.  
 XX  
 PD 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99WO-US29371.  
 XX  
 PR 10-DEC-1998; 98US-0210060.  
 XX  
 PA (SIGN-) SIGNAL PHARM INC.  
 PA (YISS) YISSUM RES & DEV CO.  
 XX  
 PI Manning AM, Mercurio F, Amlt S, Ben-eriah Y, Davis M, Hatzubai A;  
 PI Lavon I, Yaron A;  
 XX  
 DR WPI: 2000-431294/37.  
 DR N-PSDB; AAA51229.  
 XX  
 PT Polypeptide enhancing phosphorylated IkappaB ubiquitination useful for  
 PT treating disorder associated with NF-kappaB activation e.g. cancer,  
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
 PT variant  
 XX  
 PS Claim 21; Page 72-74; 77pp; English.  
 XX  
 CC Human beta-TrCP, an F-box/WD protein family member, has been shown to  
 CC have homology to human E3 ubiquitin ligase (E3). E3 enhances  
 CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of  
 CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B  
 CC degradation via the ubiquitin pathway is useful for identifying  
 CC modulators of this process for use in treating diseases associated with  
 CC activation of NF-kappa-B. In vitro analysis suggests that deletion of  
 CC the F-box results in a protein that functions as a dominant negative  
 CC molecule in vivo. Transient over-expression of endogenous I-kappa-B-alpha  
 CC deletion mutant inhibited the degradation of endogenous I-kappa-B-alpha  
 CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated  
 CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B  
 CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat  
 CC inflammatory diseases, autoimmune diseases, cancer and viral infections.  
 XX  
 SQ Sequence 569 AA;  
 XX  
 Query Match 100.0%; Score 152; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VEHSGRVRLQFDFEFOIVSSHDDTILW	29	
Db	516	VEHSGRVRLQFDFEFOIVSSHDDTILW	544	
RESULT 6				
ID	AAAY83041	standard; Protein; 569 AA.		
XX	AAAY83041			
AC	AAAY83041			
XX	16-AUG-2000	(first entry)		
DT				
XX				
DE	F-box protein FBP-1.			
XX				
KW	F-box protein; FBP; diagnosis; treatment; screening; agonist;			
KM	antagonist; proliferative disorder; differentiative disorder;			
KW	breast cancer; prostate cancer; ovarian cancer; cancer;			
KM	small cell lung carcinoma; immune disorder; cardiovascular disorder;			
XX	inflammatory disorder; human.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200012679-A1.			
XX				
PD	09-MAR-2000.			
XX				
PF	27-AUG-1999; 99WO-US19560.			
XX				
XX				
PR	28-AUG-1998; 98US-0096355.			
XX				
PR	03-FEB-1999; 99US-0118568.			
XX				
PR	15-MAR-1999; 99US-0124449.			
XX				
PA	(UYNY ) UNIV NEW YORK STATE.			
XX				
PI	Chiaur DS, Pagano M, Latres E;			
XX				
DR	WPI; 2000-256635/22.			
XX				
DR	N-PSDB; AAZ93350.			
XX				
PT	Novel nucleic acid for screening compounds useful for treating			
XX	proliferative and differentiative disorders such as cancer and immune			
PT	disorders comprises sequences encoding ubiquitin ligases			
XX				
PS	Disclosure; Figure 3a; 245pp; English.			
XX				
CC	Nucleic acids encoding substrate-targeting subunits of ubiquitin			
CC	ligases with F-box motifs (F-box proteins) are useful for diagnosis			
CC	of proliferative and differentiated related disorders by measuring			
CC	FBP gene expression. Cells expressing such proteins or			
CC	their fragments are useful for screening compounds. The compounds			
CC	are agonists or antagonists, which are useful for treating a			
CC	proliferative or differentiative disorder in a mammal such as			
CC	breast, ovarian and prostate cancer and small cell lung carcinoma			
CC	and also major opportunistic infections, immune disorders,			
CC	cardiovascular diseases and inflammatory disorders. FBP protein,			
CC	analog, derivatives and their subsequences, anti-FBP antibodies			
CC	are also useful in diagnosis of the disorders.			
XX				
SO	Sequence 569 AA;			
Query Match 100.0%; Score 152; DB 21; Length 569;				
Best Local Similarity 100.0%; Pred. NO. 6.7e-15;				
Matches	29;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	VEHSGRVRLQFDFEFOIVSSHDDTILW	29	
Db	516	VEHSGRVRLQFDFEFOIVSSHDDTILW	544	
RESULT 7				
LAY83250				

ID	AAV83250 standard; Protein; 569 AA.
XX	
AC	AAV83250;
XX	
DT	16-AUG-2000 (first entry)
XX	
DE	F-box protein hbetaRCP.
XX	
KW	Ubiquitin ligase; SCF; F-box protein; targeted degradation;
KW	destabilization; proteolysis; drug discovery; gene therapy; cancer;
KW	oncoprotein; Huntington's disease; gene knockout; delivery systems;
KW	human.
XX	
OS	Homo sapiens.
XX	
PN	WO200022110-A2.
XX	
PD	20-APR-2000.
XX	
PF	08-OCT-1999; 99WO-US23705.
XX	
PR	09-OCT-1998; 98US-0103787.
XX	
PA	(HARD ) HARVARD COLLEGE.
XX	
PT	Zhou P., Howley P;
DR	WPI; 2000-317970/27.
DR	N-PSDB; AA293710.
XX	
PT	Targeting degradation of polypeptide useful for treating cancer and
PT	other proliferative disorders, involves conjugating polypeptide with
PT	ubiquitin protein ligase or inhibiting ubiquitination using organic
XX	compound
PS	
PS	Claim 9; Page 171; 185pp; English.
XX	
CC	The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
CC	ligases) which can be used for the targeted degradation of a target
CC	polypeptide in vivo. Targeted degradation is achieved by expressing
CC	the ubiquitin ligase in a cell linked to the interaction domain of
CC	the target polypeptide and thereby recruiting the target polypeptide
CC	to the ubiquitin ligase. Such methods are useful for decreasing or
CC	increasing the level of a target polypeptide and for creating and
CC	expressing a destabilized polypeptide which is subjected to SCF
CC	mediated proteolysis. Degrading any desired protein in a cell is
CC	useful for preventing or treating diseases caused by the presence of
CC	abnormal amount of the specific polypeptides, for drug discovery and
CC	for gene therapy. Diseases treated include cancer, by degradation of
CC	oncoproteins, Huntington's disease, other proliferative disorders and
CC	microbial infections. The method provides a quick and easy
CC	alternative to gene knockout technology. The target polypeptide can
CC	be degraded at all stages, or a specific stage, of development in the
CC	mature animal.
XX	
SQ	Sequence 569 AA:
XX	
Query Match	100.0%; Score 152; DR 21; Length 569;
Best Local Similarity	100.0%; Pred. No. 6.7e-15;
Matches 29; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 VEHSGVFRLOQDEFQIVSSSHDDPTLLW 29
DB	#####
516	VEHSGVFRLOQDEFQIVSSSHDDPTLLW 544
RESULT 8	
AAV83254	
ID	AAV83254 standard; Protein; 569 AA.
XX	
AC	AAV83254;
XX	
DT	16-AUG-2000 (first entry)



DR N-PSDB; AAZ29233.  
XX  
XX Human cell signaling proteins useful for, e.g. diagnosing cell  
PT proliferative and inflammatory disorders  
PS Claim 1; Page 77-78; 90pp; English.  
XX  
XX The present sequence is cell signalling protein-12 (CSIGP-12) encoded  
CC by cDNA obtained from Incyte clone 3239149 of COLAUCP01 library. It is  
CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is  
CC found to be homologous to beta-transducin repeats containing  
CC protein. Fragments of CSIGP encoding nucleic acid can be used as  
CC hybridisation probe for detecting CSIGP related sequences or allelic  
CC variants. Recombinant CSIGP can be produced in host cells by transforming  
CC them with genetically engineered vectors. Agonists or antagonists can be  
CC used in the treatment of cell proliferative and inflammatory disorders  
CC associated with decreased or increased CSIGP expression. CSIGP is used in  
CC the diagnosis, prevention and treatment of cell proliferative disorders  
CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.  
XX  
SQ Sequence 569 AA;

Query Match 100.0%; Score 152; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 6.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVRLQFDEFOIVSSSHDDTILW 29  
|  
516 VEHSGRVRLQFDEFOIVSSSHDDTILW 544

RESULT 10  
AAB48298

ID AAB48298 standard; protein; 569 AA.

AC AAB48298;

DT 02-APR-2001 (first entry)

XX Human ZP11 protein.

XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;  
KM CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
KM Bad; Bcl-2; tumour; cytosolic.

XX Homo sapiens.

PN W0200075184-A1.

PD 14-DEC-2000.

PF 05-JUN-2000; 2000WO-US15449.

PR 04-JUN-1999; 99US-0137494.

PA (UYVA ) UNTV YALE.

PI Zhang H, Tsvetkov LM, Kondo T;

DR WPI; 2001-061703/07.

DR N-PSDB; AAC84610.

PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
involves altering levels of proteins such as S-phase kinase associated  
proteins 1, 2 and cullin/CDC53 proteins -

PS Claim 3; Page 130-132; 162pp; English.

XX The invention relates to methods of altering the polypeptide levels in a  
cell, using proteins selected from S-phase kinase associated proteins 1  
and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
cullin/CDC53 family of proteins). The method is useful for altering the

CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
CC polypeptide in a cell, SKP2 and SKP2-like protein levels are useful for  
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
CC that modulate interactions between SKP and target proteins are useful for  
CC treating tumours.

SQ Sequence 569 AA;

Query Match 100.0%; Score 152; DB 22; Length 569;  
Best Local Similarity 100.0%; Pred. No. 6.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVRLQFDEFOIVSSSHDDTILW 29  
|  
516 VEHSGRVRLQFDEFOIVSSSHDDTILW 544

RESULT 11

AAM78583

ID AAM78583 standard; Protein; 579 AA.

AC AAM78583;

DT 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1245.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KM tissue growth factor; immunomodulatory; cancer; leukaemia;  
KM nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

PN W0200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US04098.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AD, Yang Y, Wejhrman T, Goodrich R;

DR WPI; 2001-476283/51.

DR N-PSDB; AAK51716.

PT Nucleic acids encoding polypeptides with cytokine-like activities,  
useful in diagnosis and gene therapy -

XX Claim 20; Page 3504-3505; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and



CC Inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAK00020) are omitted as the relevant pages from the sequence listing CC were missing at the time of publication.

XX Sequence 579 AA:

Query Match 100.0%; Score 152; DB 22; Length 579;  
Best Local Similarity 100.0%; Pred. No. 6.8e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVRLQFDEPQIVSSSHDITLW 29  
|||||  
DB 526 VEHSGRVRLQFDEPQIVSSSHDITLW 554

RESULT 12  
AAM00847  
ID AAM00847 standard; Protein; 590 AA.

XX AAM00847;

DT 01-OCT-2001 (first entry)

DE Human bone marrow protein, SEQ ID NO: 210.

XX Human: bone marrow; antiinflammatory; cytosolic; neuroprotective;  
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;  
KW immunosuppressive; gene therapy; cytokine cell proliferation;  
KW cell differentiation modulator; immune disorder; infection; cancer;  
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.  
XX  
OS Homo sapiens.

PN WO200153453-A2.

PD 26-JUL-2001.

PF 23-DEC-2000; 2000WO-US34960.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 30-NOV-2000; 2000US-0250583.

PA (HYSE-) HYSEQ INC.

PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;

PI Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Drmanac RT;

DR WPI: 2001-488707/53.

DR N-PSDB: AAH89966.

PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful

PT for treating e.g. cancer and immune deficiency disorders -

PS Claim 10; Page 354-355; 648pp; English.

CC encodes is useful for the diagnosis and/or prognosis of one  
CC or more types of cancer. The polynucleotide and polypeptide can be  
CC used as nutritional sources or supplements and in the screening of  
CC chemical compounds as potential drugs.

XX Sequence 590 AA:

Query Match 100.0%; Score 152; DB 22; Length 590;  
Best Local Similarity 100.0%; Pred. No. 6.9e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVRLQFDEPQIVSSSHDITLW 29  
|||||  
DB 537 VEHSGRVRLQFDEPQIVSSSHDITLW 565

RESULT 13  
AAM78582  
ID AAM78582 standard; Protein; 605 AA.

XX AAM78582;

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1244.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX  
OS Homo sapiens.

PN WO200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US04098.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0596075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663351.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wehrman T, Goodrich R;

DR WPI: 2001-476283/51.

DR N-PSDB: AAK51715.

PT Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

PS Claim 20; Page 3503-3504; 6221pp; English.

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.

XX Sequence 605 AA;

Query Match 100.0%; Score 152; DB 22; Length 605;  
Best Local Similarity 100.0%; Pred. No. 7.2e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVFRLQDFEQIVSSSHDDTLILW 29  
Db 552 VEHSGRVFRLQDFEQIVSSSHDDTLILW 580

RESULT 14  
AAM00960  
ID AAM00960 standard; Protein: 608 AA.

XX AAM00960;

DT 01-OCT-2001 (first entry)

DE Human bone marrow protein, SEQ ID NO: 436.

KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;  
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;  
KW immunosuppressive; gene therapy; cytokine cell proliferation;  
KW cell differentiation modulator; immune disorder; infection; cancer;  
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.

XX Homo sapiens.

OS WO200153453-A2.

XX 26-JUL-2001.

XX 23-DEC-2000; 2000WO-US34960.

XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX 09-JUL-2000; 2000US-0598042.  
XX 19-JUL-2000; 2000US-0620312.  
XX 03-AUG-2000; 2000US-0653450.  
XX 14-SEP-2000; 2000US-0662191.  
XX 19-OCT-2000; 2000US-0693036.  
XX 30-NOV-2000; 2000US-0250583.

XX (HYSE-) HYSEQ INC.

XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Drmanac RT;

XX WPI: 2001-488707/53.

XX N-PSDB: AAM90079.

XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
PT for treating e.g. cancer and immune deficiency disorders -

XX Claim 10; Page 523-524; 648pp; English.

XX The present sequence is one of 251 novel human polypeptides encoded  
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the  
CC polypeptide encoded by it are useful in the treatment of various  
CC immune deficiencies and disorders. The deficiencies and disorders may  
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal  
CC infection, or may result from an autoimmune disorder, a coagulation  
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
CC suppression of an inflammatory response or treatment of a nervous  
CC system disorder such as Alzheimer's disease. Detection of the presence  
CC or increased expression of the polynucleotide or the protein it  
CC encodes is useful for the diagnosis and/or prognosis of one

CC or more types of cancer. The polynucleotide and polypeptide can be  
CC used as nutritional sources or supplements and in the screening of  
CC chemical compounds as potential drugs.

XX Sequence 608 AA;

Query Match 100.0%; Score 152; DB 22; Length 608;  
Best Local Similarity 100.0%; Pred. No. 7.2e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVFRLQDFEQIVSSSHDDTLILW 29  
Db 555 VEHSGRVFRLQDFEQIVSSSHDDTLILW 583

RESULT 15  
AAM78584  
ID AAM78584 standard; Protein: 632 AA.

XX AAM78584;

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1246.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

OS WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.  
XX 27-APR-2000; 2000US-0560875.  
XX 20-JUN-2000; 2000US-0598075.  
XX 19-JUL-2000; 2000US-0620325.  
XX 01-SEP-2000; 2000US-0654936.  
XX 15-SEP-2000; 2000US-0663561.  
XX 20-OCT-2000; 2000US-0693325.  
XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI: 2001-476283/51.

XX N-PSDB: AAK51717.

XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -

XX Claim 20; Page 3505-3507; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78582-AAK80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibit activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

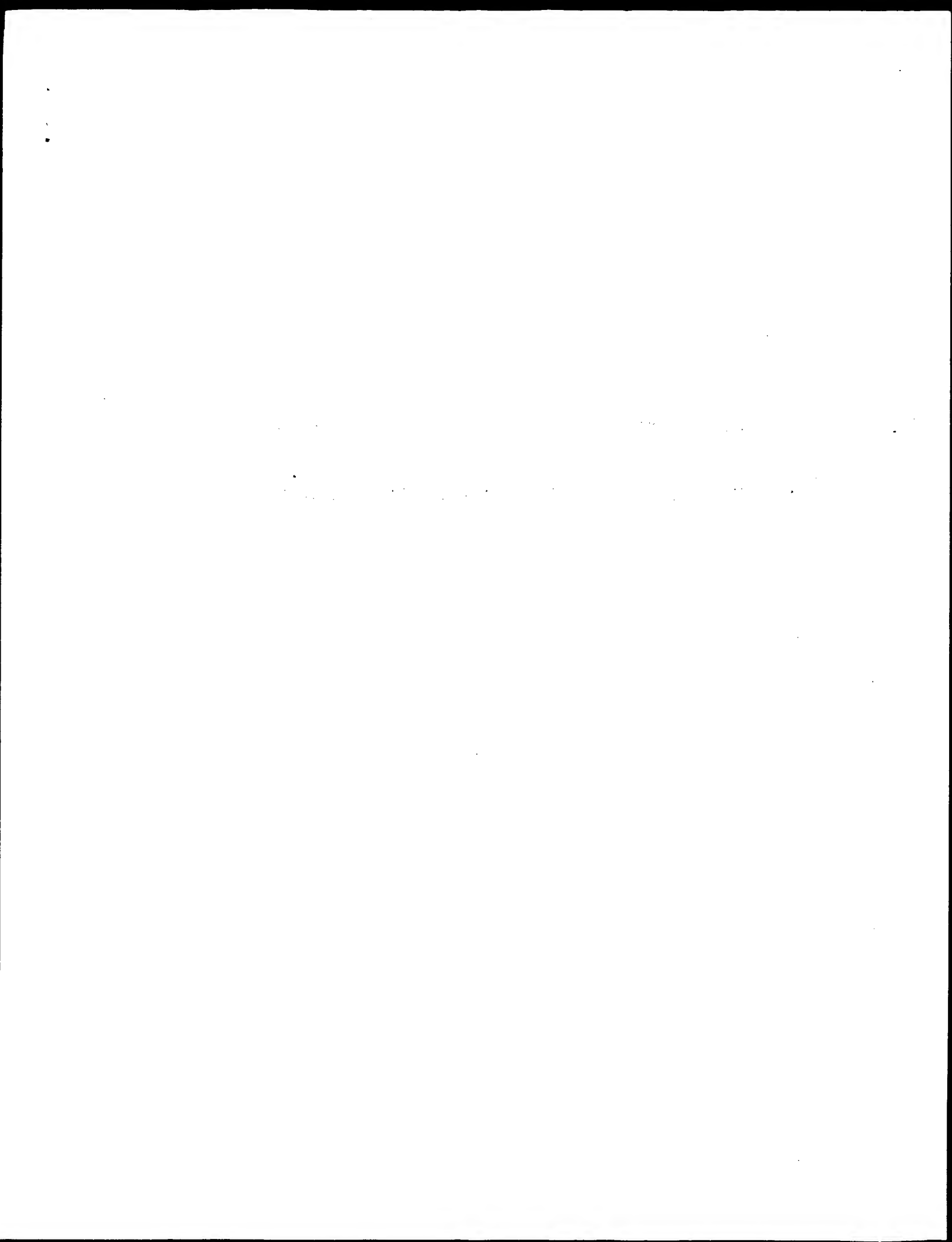
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX  
 SQ Sequence 632 AA;

Query Match 100.0%; Score 152; DB 22; Length 632;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVFRLQFDEQIVSSHDDTLIW 29  
 ||||||||||||||||||  
 Db 579 VEHSGRVFRLQFDEQIVSSHDDTLIW 607

Search completed: April 11, 2003, 11:48:17  
 Job time : 8.93666 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 10, 2003, 12:41:55 ; Search time 3.00244 Seconds

(without alignments)  
284,191 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_516\_544

Perfect score: 152  
Sequence: 1 VEHSGRVFLQFDFQIVSSSHDITLW 29Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	100.0	517	1 US-08-190-802A-30	Sequence 30, Appl
2	152	100.0	517	4 US-08-477-346-30	Sequence 30, Appl
3	152	100.0	517	4 US-08-473-089-30	Sequence 30, Appl
4	152	100.0	517	4 US-08-487-072A-30	Sequence 30, Appl
5	148	97.4	29	1 US-08-190-802A-89	Sequence 89, Appl
6	148	97.4	29	4 US-08-477-346-89	Sequence 89, Appl
7	148	97.4	29	4 US-08-473-089-89	Sequence 89, Appl
8	148	97.4	29	4 US-08-487-072A-89	Sequence 89, Appl
9	65.5	43.1	640	4 US-09-177-165A-30	Sequence 89, Appl
10	64	42.1	34	1 US-08-190-802A-83	Sequence 30, Appl
11	64	42.1	34	4 US-08-477-346-83	Sequence 83, Appl
12	64	42.1	34	4 US-08-473-089-83	Sequence 83, Appl
13	64	42.1	34	4 US-08-487-072A-83	Sequence 83, Appl
14	64	42.1	34	4 US-08-988-556A-21	Sequence 21, Appl
15	60.5	39.8	28	1 US-08-190-802A-84	Sequence 84, Appl
16	60.5	39.8	28	4 US-08-477-346-84	Sequence 84, Appl
17	60.5	39.8	28	4 US-08-473-089-84	Sequence 84, Appl
18	60.5	39.8	28	4 US-08-487-072A-84	Sequence 84, Appl
19	56	36.8	250	4 US-09-291-170A-12	Sequence 12, Appl
20	56	36.8	250	4 US-09-724-884-12	Sequence 12, Appl
21	56	36.8	704	1 US-08-188-582-18	Sequence 18, Appl
22	56	36.8	704	1 US-08-646-715-18	Sequence 18, Appl
23	56	36.8	798	1 US-08-190-802A-64	Sequence 64, Appl
24	56	36.8	798	1 US-08-190-802A-64	Sequence 64, Appl
25	56	36.8	798	2 US-08-308-818-2	Sequence 2, Appl
26	56	36.8	798	4 US-08-477-346-64	Sequence 64, Appl
27	56	36.8	798	4 US-08-477-346-64	Sequence 68, Appl

28	56	36.8	798	4 US-08-473-089-64	Sequence 64, Appl
29	56	36.8	798	4 US-08-473-089-68	Sequence 68, Appl
30	56	36.8	798	4 US-08-487-072A-64	Sequence 64, Appl
31	56	36.8	798	4 US-08-487-072A-68	Sequence 68, Appl
32	56	36.8	1194	4 US-09-092-508-2	Sequence 2, Appl
33	56	36.8	1194	4 US-09-435-115-2	Sequence 2, Appl
34	56	36.8	1194	4 US-09-069-023-26	Sequence 26, Appl
35	56	36.8	1194	4 US-09-098-310-2	Sequence 2, Appl
36	56	36.8	1205	4 US-09-092-508-16	Sequence 16, Appl
37	56	36.8	1205	4 US-09-435-115-16	Sequence 16, Appl
38	55	36.2	29	1 US-08-190-802A-86	Sequence 86, Appl
39	55	36.2	29	4 US-08-477-346-86	Sequence 86, Appl
40	55	36.2	29	4 US-08-473-089-86	Sequence 86, Appl
41	55	36.2	29	4 US-08-487-072A-86	Sequence 86, Appl
42	54	35.5	31	1 US-08-190-802A-239	Sequence 239, App
43	54	35.5	31	4 US-08-477-346-239	Sequence 239, App
44	54	35.5	31	4 US-08-473-089-239	Sequence 239, App
45	54	35.5	31	4 US-08-487-072A-239	Sequence 239, App

## ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
Sequence 30, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theroot  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ. ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-190-802A-30  
Query Match 100.0%; Score 152; DB 1; Length 517;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 VEHSGRVFLQFDFQIVSSSHDITLW 29

|||||  
DB 480 VEHSGRVRLQFDEFOIVSSHDDTILW 508

RESULT 2  
US-08-477-346-30  
; Sequence 30, Application US/08477346  
; Patent No. 6262023  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,346  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,072  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 517 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-477-346-30  
Query Match 100.0%; Score 152; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VEHSGRVRLQFDEFOIVSSHDDTILW 29  
|||||  
DB 480 VEHSGRVRLQFDEFOIVSSHDDTILW 508  
RESULT 3  
US-08-473-089-30  
; Sequence 30, Application US/08473089  
; Patent No. 6342368  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,072A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-473-089-30  
Query Match 100.0%; Score 152; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VEHSGRVRLQFDEFOIVSSHDDTILW 29  
|||||  
DB 480 VEHSGRVRLQFDEFOIVSSHDDTILW 508  
RESULT 4  
US-08-487-072A-30  
; Sequence 30, Application US/08487072A  
; Patent No. 6423684  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,072A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-487-072A-30

Query Match 100.0%; Score 152; DB 4; Length 517;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VHSGRVFRLOPDEFOIVSSSHDITLIW 29  
Db 480 VHSGRVFRLOPDEFOIVSSSHDITLIW 508

RESULT 5  
US-08-190-802A-89  
Sequence 89, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dellinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190.802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVII, Fig. 13  
US-08-190-802A-89

Query Match 97.4%; Score 148; DB 1; Length 29;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EHSGRVFRLOPDEFOIVSSSHDITLIW 29

Db 1 EHSGRVFRLOPDEFOIVSSSHDITLIW 28

RESULT 6  
US-08-477-346-89  
Sequence 89, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477.346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVII, Fig. 13  
US-08-477-346-89

Query Match 97.4%; Score 148; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EHSGRVFRLOPDEFOIVSSSHDITLIW 29  
Db 1 EHSGRVFRLOPDEFOIVSSSHDITLIW 28

RESULT 7  
US-08-473-089-89  
Sequence 89, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVII, Fig. 13  
US-08-473-089-89

Query Match  
Best Local Similarity 97.4%; Score 148; DB 4; Length 29;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EHSGRVRLQDFEQIVSSSHDITLIW 29  
DB 1 EHSGRVRLQDFEQIVSSSHDITLIW 28

RESULT 8  
US-08-487-072A-89  
Sequence 89, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Moritsen & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rVII, Fig. 13  
US-08-487-072A-89

Query Match  
Best Local Similarity 97.4%; Score 148; DB 4; Length 29;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EHSGRVRLQDFEQIVSSSHDITLIW 29  
DB 1 EHSGRVRLQDFEQIVSSSHDITLIW 28

RESULT 9  
US-09-177-165A-30  
Sequence 30, Application US/09177165A  
Patent No. 6426205  
GENERAL INFORMATION:  
APPLICANT: Tyers, Mike  
APPLICANT: Williams, Andrew  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN  
TITLE OF INVENTION: DEPENDENT PROTEOLYSIS  
FILE REFERENCE: 11757.10USU1  
CURRENT APPLICATION NUMBER: US/09/177,165A  
CURRENT FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: 60/092,443  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/063,254  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 30  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-177-165A-30

Query Match  
Best Local Similarity 43.1%; Score 65.5; DB 4; Length 640;  
Matches 13; Conservative 6; Mismatches 7; Indels 11; Gaps 1;

QY 4 SGRVRLQF-----DFEQIVSSSHDITLIW 29  
DB 541 TCKCIRIQEGHGVWDIAADNFRILSSSHGSIKRW 577

RESULT 10  
US-08-190-802A-83  
Sequence 83, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190, 802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 86000-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-190-802A-83

Query Match 42.1%; Score 64; DB 1; Length 34;  
Best Local Similarity 50.0%; Pred. No. 0.0025;  
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 EHSGRVRLQDFEFOIYSSSHDITLIIW 29  
| | | | | : | | | | | : | | | | |  
Db 6 ETSKGVYCLQYDDOKIVSGLRDNTIKIW 33

RESULT 11  
US-08-477-346-83  
Sequence 83, Application US/08477346  
Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477, 346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-477-346-83

Query Match 42.1%; Score 64; DB 4; Length 34;  
Best Local Similarity 50.0%; Pred. No. 0.0025;  
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 EHSGRVRLQDFEFOIYSSSHDITLIIW 29  
| | | | | : | | | | | : | | | | |  
Db 6 ETSKGVYCLQYDDOKIVSGLRDNTIKIW 33

RESULT 12  
US-08-473-089-83  
Sequence 83, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP rI, Fig. 13  
US-08-473-089-83

Query Match 42.1%; Score 64; DB 4; Length 34;  
Best Local Similarity 50.0%; Pred. No. 0.0025;  
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 EHSGRVRLQDFEFOIYSSSHDITLIIW 29  
| | | | | : | | | | | : | | | | |  
Db 6 ETSKGVYCLQYDDOKIVSGLRDNTIKIW 33

```

Query Match          42.1%; Score 64; DB 4; Length 43;
Best Local Similarity 40.0%; Pred. No. 0.0033;
Matches 12; Conservative 7; Mismatches 9; Indels 2; Gaps 1;

OY      2  EHSGRVRLQF--DEFOIVSSHDDTILIIW 29
      11:| | | | | | | | | | | | | | | |
Db      13  EHTGEGVGLSKSDGLQLASGGNDNTVMIIW 42

RESULT 15
US-08-190-802A-84
; Sequence 84, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived peptides and uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fadian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0860
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BETA TRCP rII, Fig. 13
US-08-190-802A-84

Query Match          39.8%; Score 60.5; DB 1; Length 28;
Best Local Similarity 44.4%; Pred. No. 0.0069;
Matches 12; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

OY      3  HSGRVRRLQFDEFOIVSSHDDTILIIW 29
      11:| | | | | | | | | | | | | | | |
Db      2  HTGSLVLCIAGYDERIVITGS-DSTVRKW 27

Search completed: April 10, 2003, 12:51:20
Job time : 4.00244 secs

```

1000 JOURNAL OF CLIMATE

FILE REFERENCE: 5914-090-999  
CURRENT APPLICATION NUMBER: US/10/042,417  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-042-417-2

Query Match 100.0%; Score 152; DB 12; Length 569;  
Best Local Similarity 100.0%; Pred. No. 3,9e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSGRVFRLQDFEFOIVSSHDDTILIW 29  
Db 516 VEHSGRVFRLQDFEFOIVSSHDDTILIW 544

RESULT 3  
US-09-213-888-7  
Sequence 7, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-213-888-7

Query Match 48.0%; Score 73; DB 9; Length 540;  
Best Local Similarity 48.1%; Pred. No. 0.0037;  
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVFRLQDFEFOIVSSHDDTILIW 29  
Db 373 HTNRVYSLQDFGIHVSGSLDTISIRW 399

RESULT 4  
US-09-213-888-10  
Sequence 10, Application US/09213888A  
Patent No. US20020164683A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/213,888A  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 10  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens

US-09-213-888-10

Query Match 48.0%; Score 73; DB 9; Length 540;  
Best Local Similarity 48.1%; Pred. No. 0.0037;  
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVFRLQDFEFOIVSSHDDTILIW 29  
Db 373 HTNRVYSLQDFGIHVSGSLDTISIRW 399

RESULT 5  
US-09-328-877A-7  
Sequence 7, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 7  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-7

Query Match 48.0%; Score 73; DB 9; Length 540;  
Best Local Similarity 48.1%; Pred. No. 0.0037;  
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVFRLQDFEFOIVSSHDDTILIW 29  
Db 373 HTNRVYSLQDFGIHVSGSLDTISIRW 399

RESULT 6  
US-09-328-877A-10  
Sequence 10, Application US/09328877A  
Patent No. US20020177187A1  
GENERAL INFORMATION:  
APPLICANT: Gurney, Mark E.  
APPLICANT: Li, Jinhe  
APPLICANT: Pauley, Adele M.  
APPLICANT: Pharmacia & Upjohn Company  
TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
FILE REFERENCE: 6142  
CURRENT APPLICATION NUMBER: US/09/328,877A  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 10  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-328-877A-10

Query Match 48.0%; Score 73; DB 9; Length 540;  
Best Local Similarity 48.1%; Pred. No. 0.0037;  
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVFRLQDFEFOIVSSHDDTILIW 29  
Db 373 HTNRVYSLQDFGIHVSGSLDTISIRW 399

```
RESULT 7
US-09-213-888-6
; Sequence 6, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-6

Query Match          48.0%; Score 73; DB 9; Length 545;
Best Local Similarity 48.1%; Pred. No. 0.0038;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVRLQDFEQIVSSSHDITLIW 29
I: ||: |||| :|| | | :|
Db 378 HTNRVYSLQFDGIHVVGSLDTSIRW 404

RESULT 8
US-09-328-877A-6
; Sequence 6, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-6

Query Match          48.0%; Score 73; DB 9; Length 545;
Best Local Similarity 48.1%; Pred. No. 0.0038;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVRLQDFEQIVSSSHDITLIW 29
I: ||: |||| :|| | | :|
Db 378 HTNRVYSLQFDGIHVVGSLDTSIRW 404

RESULT 9
US-09-213-888-5
; Sequence 5, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 9
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-5

Query Match          48.0%; Score 73; DB 9; Length 553;
Best Local Similarity 48.1%; Pred. No. 0.0038;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVRLQDFEQIVSSSHDITLIW 29
I: ||: |||| :|| | | :|
Db 386 HTNRVYSLQFDGIHVVGSLDTSIRW 412

RESULT 10
US-09-328-877A-5
; Sequence 5, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-5

Query Match          48.0%; Score 73; DB 9; Length 553;
Best Local Similarity 48.1%; Pred. No. 0.0038;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVRLQDFEQIVSSSHDITLIW 29
I: ||: |||| :|| | | :|
Db 386 HTNRVYSLQFDGIHVVGSLDTSIRW 412

RESULT 11
US-09-213-888-9
; Sequence 9, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 9
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-9

Query Match          48.0%; Score 73; DB 9; Length 553;
Best Local Similarity 48.1%; Pred. No. 0.0038;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVRLQDFEQIVSSSHDITLIW 29
I: ||: |||| :|| | | :|
Db 386 HTNRVYSLQFDGIHVVGSLDTSIRW 412
```

Query Match	48.0%;	Score 73;	DB 9;	Length 559;
Best Local Similarity	48.1%;	Pred. No. 0.0039;		
Matches 13;	Conservative	5;	Mismatches	9;
			Indels	0;
			Gaps	0;
Qy	3	HSGRVERLPDEFQIVSYSSHDPLILW	29	
		:     :     : : :		
Db	392	HTNNRVISLPDGIHVSGSLDITSIRW	418	

```

RESULT 12
US-09-328-877A-9
; Sequence 9, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Guiney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 9
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-9

```

Query Match	48.0%;	Score 73;	DB 9;	Length 559;
Best Local Similarity	48.1%;	Pred. No. 0.0039;		
Matches 13;	Conservative	5;	Mismatches	9;
			Indels	0;
			Gaps	0;
Qy	3	HSGRVERLQDFEQIVSSSDHDTILTW	29	
Db	392	HTNREVIQLQDFGIHVVGSGSLDTSIRFW	418	

```

RESULT 13
US-09-213-888-8
: Sequence 8 / Application US/09213888A
: Patent No. US20020164683A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Li, Jinhe
: APPLICANT: Pauley, Adele M.
: APPLICANT: Pharmacia & Upjohn Company
: TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
: FILE REFERENCE: 6142
: CURRENT APPLICATION NUMBER: US/09/213.888A
: CURRENT FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 8
: LENGTH: 589
: TYPE: PR1
: ORGANISM: Homo sapiens
US-09-213-888-8

```

Query Match	48.0%	Score 73	DB 9	Length 589
Best Local Similarity	48.1%	Pred. No. 0.0041		
Matches 13	Conservative	5	Mismatches	9
			Indels	0
			Gaps	0
QY	3	HSGRVRLQDFDEQIYSSSDITILIM	29	
Db	422	HTNRVYSLQFDGIHVVSGSLDTSIRVW	448	

RESULT 14  
US-09-328-877A-8

```

1 Sequence 8 Application US/09328877A
2 Patent No. US20020177187A1
3 GENERAL INFORMATION:
4 APPLICANT: Gurney, Mark E.
5 APPLICANT: Li, Jinhe
6 APPLICANT: Pauley, Adele M.
7 TITLE OF INVENTION: Pharmacia & Upjohn Company
8 TITLE OF INVENTION: Human Ser-10 Polypeptides and Polynucleotides that
9 FILE REFERENCE: 6142
10 CURRENT APPLICATION NUMBER: US/09/328, 877A
11 CURRENT FILING DATE: 1999-06-09
12 NUMBER OF SEQ ID NOS: 27
13 SOFTWARE: PatentIn Ver. 2.0
14 SEQ ID NO 8
15 LENGTH: 589
16 TYPE: PRT
17 ORGANISM: Homo sapiens
18 US-09-328-877A-8

```

Query Match	48.0%	Score 73	DB 9	Length 589
Best Local Similarity	48.1%	Pred. No. 0.0041		
Matches 13	Conservative 5	Mismatches 9	Indels 0	Gaps 0
Qy	3	HSGRVFRLQDFEQTIVSSSHDTILIM	29	
	: : : : : : : : : : : : : :			
Db	422	HTNKEVSLQDFGIHVSGSLDTSIRWV	448	

```

RESULT 15
US-09-213-888-4
: Sequence 4, Application US/09213888A
: Patent No. US20020164683A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Li, Jinhe
: APPLICANT: Pauley, Adele M.
: APPLICANT: Pharmacia & Upjohn Company
: TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
: FILE REFERENCE: 6142
: CURRENT APPLICATION NUMBER: US/09/213,888A
: CURRENT FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 592
: TYPE: prt
: ORGANISM: Homo sapiens
: US-09-213-888-4

```

	Query Match:	48.0%;	Score 73;	DB 9;	Length 592;
	Best Local Similarity	48.1%;	Pred. No. 0.0041;		
	Matches	13;	Conservative	5;	Mismatches 9; Indels 0; Gaps 0.
OY	3 HSGVRFLODFEQIVSSSHDITLILW	29			
	I : I I : I I I I : I I I I : I I				
Db	425 HINRNVSLQPDGLIHVVGSLSLTISRW	451			

Search completed: April 10, 2003, 13:16:47  
Job time : 2.96711 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: April 10, 2003, 13:15:21 ; Search time 3.35566 Seconds  
(without alignments)  
830.804 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_516\_544

Perfect score: 152  
Sequence: 1 VEHSGRVFRLQDFEFOIVSSSHDDTLILW 29

Scoring table: BLOSOM62  
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: PIR:73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	100.0	518	2	B48088
2	143	94.1	701	2	beta-transducin re
3	70	46.1	267	2	hypothetical prote
4	70	46.1	506	2	hypothetical trp-a
5	69	45.4	605	2	WD-repeat protein
6	66	43.4	650	2	probable sulfur me
7	65.5	43.1	640	2	sulfur controller-
8	64	42.1	610	2	MET10 protein - ye
9	62	40.8	611	2	cell division cont
10	61	40.1	651	2	hypothetical prote
11	57	37.5	330	2	WD repeat protein
12	57	37.5	370	2	hypothetical prote
13	57	37.5	377	2	GTP-binding protei
14	57	37.5	377	2	GTP-binding protei
15	57	37.5	377	2	GTP-binding protei
16	57	37.5	377	2	GTP-binding protei
17	57	37.5	377	2	GTP-binding protei
18	57	37.5	380	2	GTP-binding protei
19	57	37.5	380	2	GTP-binding protei
20	57	37.5	380	2	GTP-binding protei
21	56.5	37.2	424	2	hypothetical prote
22	56.5	36.8	411	2	G1/S transition co
23	56	36.8	316	2	guanine nucleotide
24	56	36.8	377	2	GTP-binding protei
25	56	36.8	496	2	GTP-binding protei
26	56	36.8	715	2	hypothetical prote
27	56	36.8	793	2	DOAI protein - yea
28	56	36.8	1194	2	TATA box-binding p
29	55.5	36.5	415	2	apoptotic proteina
					probable WD-40 rep

30	55.5	36.5	612	2	T08602	protein Tfpd - s11
31	55	36.2	290	2	T02300	GTP-binding regula
32	55	36.2	325	2	T23309	hypothetical prote
33	55	36.2	600	2	T23497	hypothetical prote
34	55	36.2	1683	2	AF2071	WD-40 repeat prote
35	54.5	35.9	1008	2	T32986	hypothetical prote
36	54	35.5	317	2	A33928	GTP-binding protei
37	54	35.5	317	2	B33928	GTP-binding protei
38	54	35.5	317	2	S45054	GTP-binding regula
39	54	35.5	317	2	S38398	GTP-binding regula
40	54	35.5	317	2	A36986	activated protein
41	54	35.5	318	2	S11904	GTP-binding regula
42	54	35.5	325	2	T06784	GTP-binding protei
43	54	35.5	325	2	T09613	probable GTP-bindi
44	54	35.5	327	2	S48839	guanine nucleotide
45	54	35.5	356	2	T22478	hypothetical prote

## ALIGNMENTS

RESULT 1  
B48088  
beta-transducin repeat-containing protein - African clawed frog  
N:Alternate names: beta-Trcp  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 26-May-1994 #sequence-revision 26-May-1994 #text-change 21-Jul-2000  
C:Accession: B48088  
R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.  
Mol. Cell. Biol. 13, 4953-4966, 1993  
A>Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase  
A:Reference number: A48088; MUID:933030289; PMID:8393141  
A:Accession: B48088  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-518 <SPE>  
A:Cross-references: GB:M98268; NID:9295542; PIDN:AAA02810.1; PID:9295543  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: duplication  
F:431-462/Domain: WD repeat homology <WD1>

Query Match 100.0%; Score 152; DB 2; Length 518;  
Best Local Similarity 100.0%; Pred. No. 1.3e-14;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVFRLQDFEFOIVSSSHDDTLILW 29  
DB 481 VEHSGRVFRLQDFEFOIVSSSHDDTLILW 509

## RESULT 2

T16607  
hypothetical protein K10B2.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Sep-1999 #sequence-revision 20-Sep-1999 #text-change 20-Sep-1999  
C:Accession: T16607  
R:Miller, N.  
submitted to the EMBL Data Library, June 1995  
A:Description: The sequence of C. elegans cosmid K10B2.  
A:Reference number: Z18545  
A:Accession: T16607  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-701 <MIU>  
A:Cross-references: EMBL:028730; NID:9660694; PID:9660695; PIDN:AAA68258.1; CESP:K10B  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:K10B2.1  
A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 94.1%; Score 143; DB 2; Length 701;  
Best Local Similarity 89.7%; Pred. No. 3.9e-13;  
Matches 26; Conservative 3; Mismatches 0; Indels 0; Gaps 0;





```

A:Residues: 1-60; 'I','62'-640 <THO>
A:Cross-references: EMBL:L26505; NID:g432493; PID:g432494
C:Genetics:
A:Gene: SGD:MET30
A:Cross-references: SGD:S0001308; MIPS:YIL046w
A:Map position: 9L
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
F:298-329/Domain: WD repeat homology <WD1>
F:338-369/Domain: WD repeat homology <WD2>
F:374-409/Domain: WD repeat homology <WD3>
F:417-450/Domain: WD repeat homology <WD4>
Query Match          43.1%; Score 65.5; DB 2; Length 640;
Best Local Similarity 35.1%; Pred. No. 0.12;
Matches 13; Conservative 6; Mismatches 7; Indels 11; Gaps 1;

QY      4 SGRVFRLOF-----DEFOIVSSHDDITLIW 29
       :|:| | | | | | | | | | | | | | | | |
Db      541 TKGCIPTQFGHEGVWDIAADNFRITISGSHDSIKRW 577

RESULT 8
S64126
cell division control protein CDC20 - yeast (Saccharomyces cerevisiae)
N:Alternate names: Protein G2939; Protein YG116w
C:Species: Saccharomyces cerevisiae
C:date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 26-May-2000
C:Accession: S64126; S18840; S48507
R:Ianguin, G.
Submitted to the Protein Sequence Database, May 1996
A:Reference number: S64122
A:Accession: S64126
A:Molecule type: DNA
A:Residues: 1-610 <LANU>
A:Cross-references: EMBL:L272638; NID:g1322668; PIDD:CAA96824.1; PID:e243344; PID:g1322668
A:Experimental source: strain S288C
R:Sechi, N.; Montaguado, M.C.; Koshland, D.; Hogan, E.; Burke, D.J.
Mol. Cell. Biol. 11, 5592-5602, 1991
A>Title: The CDC20 gene product of Saccharomyces cerevisiae, a beta-transducin homolog,
A:Reference number: S18840; MUID:92017840; PMID:1922065
A:Accession: S18840
A:Molecule type: DNA
A:Residues: 1-317; 'MA','320-501','ORMOSLFITIKONS','514','1OK' <SET>
A:Cross-references: EMBL:X59428; NID:g3644; PIDD:CAA42058.1; PID:g3645
A>Note: The authors translated the codon TAT for residue 220 as Gln
R:Dol, A.; Dol, K.
submitted to the EMBL Data Library, June 1993
A>Description: Correct end of the ORF for the CDC20 gene of Saccharomyces cerevisiae.
A:Reference number: S48507
A:Accession: S48507
A:Molecule type: DNA
A:Residues: 50-610 <DO1>
A:Cross-references: EMBL:D16506; NID:g391938; PIDD:BAA03957.1; PID:d1004473; PID:g416288
C:Comment: This is a G-beta protein.
C:Genetics:
A:Gene: SGD:CDC20
A:Cross-references: SGD:S0003084; MIPS:YGL16w
A:Map position: 7L
C:function:
A:description: required for nuclear movements prior to anaphase; required for chromosome
C:superfamily: unassigned WD repeat proteins; WD repeat homology
F:381-414/Domain: WD repeat homology <WD1>

Query Match          42.1%; Score 64; DB 2; Length 610;
Best Local Similarity 40.0%; Pred. No. 0.19;
Matches 12; Conservative 7; Mismatches 9; Indels 2; Gaps 1;

QY      2 EHSGRVFRLOF--DEFOIVSSHDDITLIW 29
       ||||| | | | | | | | | | | | | | | | |
Db      383 EHNGEVCGLSYKSDGIQLASCGMDNTVMIM 412

RESULT 9
```

```

T15410      hypothetical protein C04F6.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 26-May-2000
C:Accession: T15410
R:Nhan, M.
Submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid C04F6.
A:Reference number: 218346
A:Accession: T15410
A:Status: preliminary; translated from GB/EMBL/DDbJ
A:Molecule type: DNA
A:Residues: 1-611 <NHA>
A:Cross-references: EMBL:U42835; NID:g1125760; PID:g1125764; PIDs:AAA83588.1; CESP:C0
C:Genetics:
A:Gene: CESP:C04F6.4
A:Introns: 25/1; 184/3; 349/1; 514/3; 577/1
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match          40.8%; Score 62; DB 2; Length 611;
Best Local Similarity 51.6%; Pred. No. 0.37;
Matches 16; Conservative 3; Mismatches 10; Indels 2; Gaps 1;

QY       1 VEHSGRVFRLOF--DEFOIVSSSHDITLIW 29
           ||||| | | : | : | | | | | |
Db        236 VAHSGSVFGLTWSPDCTKIASASADKTIKIW 266

RESULT 10
T50289
WD repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50289
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25054
A:Accession: T50289
A>Status: preliminary; translated from GB/EMBL/DDbJ
A:Molecule type: DNA
A:Residues: 1-651 <WOO>
A:Cross-references: EMBL:AL157993; PTDN:CAB76232.1; GSPDB:GN00066; SPDB:SPACUNK12.01
A:Experimental source: strain 972h(-); cosmid c25D11
C:Genetics:
A:Gene: SPDB:SPACUNK12.01
A:Map position: 1

Query Match          40.1%; Score 61; DB 2; Length 651;
Best Local Similarity 34.5%; Pred. No. 0.56;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY       1 VEHSGRVFRLOFDEFOIVSSSHDITLIW 29
           ||| : |||||
Db         533 IGHTAPISSLQFDNSHLHLSGSTDNSYRIW 561

RESULT 11
S70127
hypothetical protein YDR267c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D9954.11
C:Species: Saccharomyces cerevisiae
C:Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 19-Apr-2002
C:Accession: S70127
R:Ile, T.
submitted to the EMBL Data Library, May 1996
A:Description: The sequence of S. cerevisiae cosmid 9954.
A:Reference number: S70124
A:Accession: S70127
A:Molecule type: DNA
A:Residues: 1-330 <LEV>
A:Cross-references: EMBL:U51030; NID:g1332633; PID:g1230640; GSPDB:GN00004; MIPS:YDR267c
A:Gene: MIPS::YDR267c

```

A:Cross-references: SGD:S0002675

A:Map position: 4R

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

F:54-87/Domain: WD repeat homology <WD1>

F:103-136/Domain: WD repeat homology <WD2>

F:149-182/Domain: WD repeat homology <WD3>

Query Match 37.5%; Score 57; DB 2; Length 330;

Best Local Similarity 35.5%; Pred. No. 1;

Matches 11; Conservative 6; Mismatches 10; Indels 4; Gaps 1;

OY 3 HSGRVFRLQF---DEFOIVSSSHDDITLILW 29  
 Db 196 HEGTVWSDPDKTEGVFRLCSGSDSTVRVW 226

#### RESULT 12

T03256

GTP-binding protein beta chain - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 20-Jun-2000

C:Accession: T03256

R:Oelmeier, R.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z14864

A:Accession: T03256

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-375 <OEI>

A:Cross-references: EMBL:X8161; NID:g1360091; PIDN:CA66842.1

A:Experimental source: cultivar Samsun NN; tissue-type leaf

C:Superfamily: GTP-binding regulatory protein beta chain; WD repeat homology

C:Keywords: GTP binding; heterotrimer; signal transduction

F:244-277/Domain: WD repeat homology <WDR>

Query Match 37.5%; Score 57; DB 2; Length 375;

Best Local Similarity 34.5%; Pred. No. 1.2;

Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 1;

OY 3 HSGRVFRLQF---DEFOIVSSSHDDITLILW 29

Db 64 HTGKVISLDMWTPKRNRIVSASODGRLIYW 92

#### RESULT 13

T16985

GTP-binding protein beta chain - curled-leaved tobacco

N:Alternate names: G protein beta chain

C:Species: Nicotiana glauca (curled-leaved tobacco)

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Jun-2000

C:Accession: T16985

R:Kaydanov, C.; Tewes, A.; Mantouf, R.

submitted to the EMBL Data Library, November 1996

A:Description: Molecular cloning of a G protein alpha and beta subunit from somatic embryo

A:Reference number: Z18624

A:Accession: T16985

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-377 <KAY>

A:Cross-references: EMBL:Y09513

A:Experimental source: somatic embryo

C:Superfamily: GTP-binding regulatory protein beta chain; WD repeat homology

C:Keywords: GTP binding; signal transduction

F:244-277/Domain: WD repeat homology <WDR>

Query Match 37.5%; Score 57; DB 2; Length 377;

Best Local Similarity 34.5%; Pred. No. 1.2;

Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 1;

OY 3 HSGRVFRLQF---DEFOIVSSSHDDITLILW 29

Db 64 HTGKVISLDMWTPKRNRIVSASODGRLIYW 92

#### RESULT 14

T07376

G-protein beta chain Gb1 - potato

N:Alternate names: Gb1 protein

C:Species: Solanum tuberosum (potato)

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Jun-2000

C:Accession: T07376

R:Provart, N.J.

submitted to the EMBL Data Library, June 1995

A:Reference number: Z15997

A:Accession: T07376

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-377 <PRO>

A:Cross-references: EMBL:X87837

A:Experimental source: cv. Desiree; dev. stage 8 weeks; tissue type epidermal fragment

C:Genetics:

A:Note: gbl

C:Complex: heterotrimer

C:Function:

A:Description: belongs to a family of guanine nucleotide-binding proteins that relay

C:Superfamily: GTP-binding regulatory protein beta chain; WD repeat homology

C:Keywords: GTP binding; heterotrimer; signal transduction

F:244-277/Domain: WD repeat homology <WDR>

Query Match 37.5%; Score 57; DB 2; Length 377;

Best Local Similarity 34.5%; Pred. No. 1.2;

Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 1;

OY 3 HSGRVFRLQF---DEFOIVSSSHDDITLILW 29

Db 64 HTGKVISLDMWTPKRNRIVSASODGRLIYW 92

#### RESULT 15

T04086

GTP binding protein beta chain - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000

C:Accession: T04086

R:Lein, W.; Saalbach, G.

submitted to the EMBL Data Library, February 1997

A:Description: Characterization of G-protein function in plants.

A:Reference number: Z15198

A:Accession: T04086

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-377 <LEI>

A:Cross-references: EMBL:Z84820; PIDN:CAB06618.1

A:Experimental source: strain SRI; tissue-type leaf

C:Superfamily: GTP-binding regulatory protein beta chain; WD repeat homology

F:244-277/Domain: WD repeat homology <WDR>

Query Match 37.5%; Score 57; DB 2; Length 377;

Best Local Similarity 34.5%; Pred. No. 1.2;

Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 1;

OY 3 HSGRVFRLQF---DEFOIVSSSHDDITLILW 29

Db 64 HTGKVISLDMWTPKRNRIVSASODGRLIYW 92

Search completed: April 10, 2003, 13:23:06

Job time : 4.35566 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 10, 2003, 12:51:26 ; Search time 2.08404 Seconds  
(without alignments)  
577.154 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_516\_544

Sequence: 1 VEHSGRVRLQFDEFQIVSSHDDTILIW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	100.0	518	TRCB_XENLA	Q91854 xenopus lae
2	152	100.0	605	FW1A_HUMAN	Q9Y297 homo sapien
3	151	99.3	542	FW1B_HUMAN	Q9UKB1 homo sapien
4	143	94.1	665	L123_CAEEL	Q09990 caenorhabd
5	70	46.1	506	POFB_SCHPO	Q09855 schizosacch
6	69	45.4	605	POFI_SCHPO	P87053 schizosacch
7	66	43.4	650	SC02_NEUCR	Q01277 neuropeptid
8	65.5	43.1	640	MT30_YEAST	P39314 saccharomyc
9	64.5	42.4	525	CO2A_HUMAN	Q92828 homo sapien
10	64	42.1	610	CC20_YEAST	P26309 saccharomyc
11	62	40.8	611	MDRI_CAEEL	Q11176 caenorhabd
12	61	40.1	651	YIXI_SCHPO	Q09713 schizosacch
13	61	40.1	678	SC0B_EMENT	Q00659 emericella
14	57	37.5	375	GBB3_TOBAC	P93397 nicotiana t
15	57	37.5	377	GBB1_TOBAC	P93398 nicotiana t
16	57	37.5	377	GBB2_TOBAC	P93399 nicotiana t
17	57	37.5	377	GBB_NICPL	P93363 solanum tub
18	57	37.5	377	GBB_SOLTU	P49178 zea mays (m
19	57	37.5	380	GBB_MAIZE	Q40667 oryza sativ
20	57	37.5	380	GBB_ORNSA	P47025 saccharomyc
21	57	37.5	714	YJL2_YEAST	Q22469 arbidopsin
22	56.5	37.2	424	MS13_NEUCR	Q02349 neuropeptid
23	56	36.8	316	GBLP_NEUCR	Q02189 hydra atten
24	56	36.8	317	GBLP_HYDAT	P49177 arbidopsin
25	56	36.8	377	GBB_ARATH	P36037 saccharomyc
26	56	36.8	715	DOAI_YEAST	P38129 saccharomyc
27	56	36.8	798	T2D4_YEAST	Q15352 homo sapien
28	56	36.8	800	APAF_HUMAN	O14727 homo sapien
29	55.5	36.8	1248	MS12_ARATH	Q22468 arbidopsin
30	55.5	36.5	612	TIPD_DICDI	O15736 dicystosell
31	55.5	36.5	2295	WR9_HUMAN	Q9NS16 homo sapien
32	55.5	36.5	324	GBLP_CAEEL	Q21215 caenorhabd
33	55	36.2	1	GBLP_CAEEL	Q21215 caenorhabd

34	55	36.2	1683	YJL24_ANASP	Q8YV57 anabaena sp
35	54	35.5	316	GBLP_BIOGL	Q93134 biophthalari
36	54	35.5	317	GBLP_BRARE	O42248 brachydanio
37	54	35.5	317	GBLP_HUMAN	P25388 homo sapien
38	54	35.5	317	GBLP_ORYZ	O42249 oreochromis
39	54	35.5	318	GBLP_ORYZ	P25387 chlamydomon
40	54	35.5	325	GBLP_MEDSA	O24076 medicago sa
41	54	35.5	325	GBLP_SOYBN	Q39836 glycine max
42	54	35.5	327	GBLP_ARATH	O24456 arbidopsin
43	54	35.5	327	GBLP_BRANA	Q39336 brassica na
44	54	35.5	356	GBB2_CAEEL	Q20636 caenorhabd
45	54	35.5	422	FBW2_HUMAN	Q9UKB8 homo sapien

## ALIGNMENTS

```

RESULT 1
ID TRCB_XENLA STANDARD: PRT; 518 AA.
AC Q91854; P70037; P70038;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-Trip (beta-transducin repeat-containing protein).
GN FBXW1 OR BTFCP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP MEDLINE=9330289; PubMed=8393141;
RX Sevak W., Keiper B.D., Stratowa C., Castanon M.J.;
"Xenopus laevis cdc15 mutants arrested at a late stage in
anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
with beta-transducin repeats."
RT Mol. Cell. Biol. 13:4953-4966(1993).
RL [2]
RN SEQUENCE OF 302-518 FROM N.A.
RP MEDLINE=97109804; PubMed=8952061;
RX Hudson J.W., Alarcon V.B., Rinson R.P.;
"Identification of new localized RNAs in the Xenopus oocyte by
RT differential display PCR."
RL Dev. Genet. 19:190-196(1996).
CC - FUNCTION: Probably recognizes and binds to some phosphorylated
proteins and promotes their ubiquitination and degradation.
CC - SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX
(cc BY SIMILARITY).
CC - DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
MAURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
TADPOLE EMBRYO.
CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC - SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC - SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: M98268; AAA02810.1; -
CC EMBL: U63922; AAA49672.1; -
CC EMBL: U63922; AAA49672.1; -
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001810; WD40.
CC Pfam: PF00400; WD40; 7.

```

DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ubl conjugation pathway; Repeat; WD repeat.  
 FT DOMAIN 119 157  
 F-BOX.  
 FT REPEAT 230 258  
 WD 1.  
 FT REPEAT 270 298  
 WD 2.  
 FT REPEAT 310 338  
 WD 3.  
 FT REPEAT 353 381  
 WD 4.  
 FT REPEAT 393 421  
 WD 5.  
 FT REPEAT 433 461  
 WD 6.  
 FT REPEAT 482 510  
 WD 7.  
 FT CONFLICT 302 304  
 GEM -> EFR (IN REF. 2).  
 FT CONFLICT 516 518  
 GEM -> AAH (IN REF. 2).  
 SQ SEQUENCE 518 AA; 59507 MW; 2A52EC19028127F3 CRC64;

Query Match 100.0%; Score 152; DB 1; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEHSGRVRLQDFEQIVSSSHDITLILW 29  
 DB 481 VEHSGRVRLQDFEQIVSSSHDITLILW 509

## RESULT 2

ID FW1A\_HUMAN STANDARD; PRT; 605 AA.  
 AC Q9Y297; Q9Y213;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-Trcp)  
 DE (E3K8IKAPPB) (pIkappabalpha-E3 receptor subunit).  
 GN FBXW1A OR FBW1A OR BTRCP OR BTRC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=99075339; PubMed=9859996;  
 RA Yaron A., Hatzubai A., Davis M., Lavon I., Manning A.M.,  
 RT Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
 RT "Identification of the receptor component of the ikappabalpha-  
 RT ubiquitin ligase.";  
 RL Nature 396:590-594(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 KC TISSUE=Lymphoid;  
 RX MEDLINE=98325370; PubMed=9660940;  
 RA Margotelin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,  
 RA Thomas D., Strebel K., Benarous R.,  
 RT "A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu  
 RT connects CD4 to the ER degradation pathway through an F-box motif.";  
 RL Mol. Cell 1:565-574(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Cenciarelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99145464; PubMed=9990852;  
 RA Winston J.T., Strack P., Beer-Komer P., Chu C.Y., Ellledge S.J.,

RA Harper J.W.;  
 RT "The SCF(beta-Trcp)-ubiquitin ligase complex associates specifically  
 RT with phosphorylated destruction motifs in I-kappa-B-alpha and  
 RT beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro.";  
 RL Genes Dev. 13:270-283(1999).  
 CC -1- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA  
 CC (PIKAPPAALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR  
 CC UBIQUITINATION AND DEGRADATION.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2, ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF101784; AAD08702.1; -  
 DR EMBL: Y14153; CAAT4572.1; -  
 DR EMBL: AF129530; AAF04464.1; -  
 DR Gene; HGNC:1144; BTRC.  
 DR MIM: 603482; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00440; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ubl conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 FT DOMAIN 190 228  
 F-BOX.  
 FT REPEAT 301 338  
 WD 1.  
 FT REPEAT 341 378  
 WD 2.  
 FT REPEAT 381 418  
 WD 3.  
 FT REPEAT 424 461  
 WD 4.  
 FT REPEAT 464 503  
 WD 5.  
 FT REPEAT 505 541  
 WD 6.  
 FT REPEAT 553 590  
 WD 7.  
 FT VARSPLIC 17 52  
 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 605 AA; 68866 MW; 4C67F3B7E40FD37 CRC64;

Query Match 100.0%; Score 152; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEHSGRVRLQDFEQIVSSSHDITLILW 29  
 DB 552 VEHSGRVRLQDFEQIVSSSHDITLILW 580

## RESULT 3

ID FW1B\_HUMAN STANDARD; PRT; 542 AA.  
 AC Q9UKB1; Q9Y4C6; Q9P2S8; Q9P2S9;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-Trcp2).  
 GN FBXW1B OR FBW1B OR BTRCP2 OR KIAA0696.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBL\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=20003060; PubMed=10531035;  
 RA Genciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins.";  
 RL Curr. Biol. 9:1177-1179(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Petal Lung;  
 RX MEDLINE=20160458; PubMed=10694485;  
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,  
 RA Katoh M.;  
 RT "Molecular cloning and genomic structure of the betatropin gene on  
 RT chromosome 5q35.1.";  
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B AND C (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF176022; AAF04528.1; -;  
 DR EMBL: AB033329; BAA92329.1; -;  
 DR EMBL: AB033380; BAA92330.1; -;  
 DR EMBL: AB033381; BAA92331.1; -;  
 DR EMBL: AB014596; BAA31671.1; ALT\_INIT.  
 DR Genew: HGNC:13607; FBXW1B.  
 DR MIM: 605651; -;  
 DR Interpro: IPR001810; F-box.  
 DR Interpro: IPR001680; WD40.  
 DR Pfam: PF00440; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00320; WD40; 7.  
 DR SMART: SM00320; WD40; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE: PS00678; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Ub1 conjugation pathway; Repeat: WD repeat; Alternative splicing.  
 FT DOMAIN 129 167 F-BOX.  
 FT REPEAT 238 275 WD 1.  
 FT REPEAT 278 315 WD 2.  
 FT REPEAT 318 355 WD 3.  
 FT REPEAT 361 398 WD 4.  
 FT REPEAT 401 440 WD 5.  
 FT REPEAT 442 478 WD 6.  
 FT REPEAT 490 527 WD 7.  
 FT VARSPLIC 16 49 MISSING (IN ISOFORM A).

FT VARSPLIC 16 48 CSVPRLMLGCAVLVESGALSLQSPVSVCL -> NMSV  
 FT SEQUENCE 542 AA: 62090 MW: 7CD40087EFAA5C8A CRC64;  
 SQ MEDONDESPKKNIL (IN ISOFORM B).  
 Query Match 99.3%; Score 151; DB 1; Length 542;  
 Best Local Similarity 96.6%; Pred. No. 2.8e-15;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VEHSGRVFLRLODFEFOIIVSSSHDITILIM 29  
 DB 489 VEHSGRVFLRLODFEFOIIVSSSHDITILIM 517  
 RESULT 4  
 ID L123 CAEFL STANDARD; PRT; 665 AA.  
 AC Q09990; Q9GNN6;  
 DT 15-JUN-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein lln-23.  
 GN L1N-23 OR K10B2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Pelodermidae; Caenorhabditis.  
 OX NCBL\_TaxID=6239;  
 RN [1]  
 RN SEQUENCE FROM N.A., FUNCTION, DEVELOPMENTAL STAGE, AND MUTAGENESIS.  
 RP MEDLINE=20515608; PubMed=11060233;  
 RA Kipreos E.T., Gohel S.P., Hedgecock E.M.;  
 RT "The Caenorhabditis elegans F-box/WD-repeat protein lln-23 functions  
 RT to limit cell division during development.";  
 RL Development 127:5071-5082(2000).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Bristol N2;  
 RA Miller N.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Functions cell autonomously to negatively regulate cell  
 CC cycle progression. Required to restrain cell proliferation in  
 CC response to developmental cues. Probably recognizes and binds to  
 CC some proteins and promotes their ubiquitination and degradation  
 CC (By similarity).  
 CC -1- SUBUNIT: Part of a SCF (SKP1-CULLIN-F-BOX) protein ligase complex  
 CC (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- DEVELOPMENTAL STAGE: Highest levels in embryos and adults, lowest  
 CC levels in larvae. Maternal expression results in high zygotic  
 CC levels.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO X.LAEVIS FBXW1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF275253; AAG28037.1; -;  
 DR EMBL: U28730; AAA68258.2; -;  
 DR WormPep: K10B2.1; CE28600.  
 DR Interpro: IPR001810; F-box.  
 DR Interpro: IPR001680; WD40.  
 DR Pfam: PF00440; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.

```

DR Prodom; PD000018; WD40; 4.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 5.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Developmental protein; Cell cycle; Cell division;
KW ubi conjugation pathway; Repeat; WD repeat.
FT DOMAIN 81 127 F-BOX.
FT REPEAT 220 257 WD 1.
FT REPEAT 260 299 WD 2.
FT REPEAT 301 337 WD 3.
FT REPEAT 343 380 WD 4.
FT REPEAT 383 420 WD 5.
FT REPEAT 423 460 WD 6.
FT REPEAT 472 509 WD 7.
FT DOMAIN 525 582 WD 7.
FT MOTIF 441 441 ALA-RICH.
SQ SEQUENCE 665 AA; 75916 MW; BF3E9AF5F12ECCC CRC64;

Query Match
Best Local Similarity 94.1%; Score 143; DB 1; Length 665;
Matches 26; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEHSGRVRLQDFDEQIVSSSHDITLIW 29
Db 471 VQHTGRVFRLODFDFIVSSSHDITLIW 499

RESULT 5
POF1_SCHPO STANDARD; PRT; 506 AA.
AC 009855; 09P7V1.
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof1.
GN POF1 OR SPAC29E6.01 OR SPAC30.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrison C.L., Toda T.;
RT "Systematic genome-wide analysis of F-box protein-encoding genes in
RT fission yeast."
RT Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris P., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth N., Huckle E.J., Hunt S., Jagels K.,
RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
RA Weltens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lemrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Weiler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

```

```

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Bartell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -I- SIMILARITY: CONTAINS 7 WD_REPEATS (TRP-ASP DOMAINS).
CC -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB061694; BAB55543.1; -
DR EMBL; AL136538; CAB66464.1; -
DR EMBL; Z66525; CA91423.1; -
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPROTEINRPT.
DR Prodom; PD000018; WD40; 3.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 8.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 3.
DR PROSITE; PS50082; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT DOMAIN 70 116 F-BOX.
FT REPEAT 219 256 WD 1.
FT REPEAT 259 298 WD 2.
FT REPEAT 301 338 WD 3.
FT REPEAT 345 386 WD 4.
FT REPEAT 388 426 WD 5.
FT REPEAT 427 464 WD 6.
FT REPEAT 468 505 WD 7.
SQ SEQUENCE 506 AA; 58257 MW; CEF34DAEFFBC2E10 CRC64;

Query Match
Best Local Similarity 46.1%; Score 70; DB 1; Length 506;
Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 4 SGRVRLQDFDEQIVSSSHDITLIW 29
Db 470 NSRVEGLQDFHRIITACTHSEILIW 495

RESULT 6
POF1_SCHPO STANDARD; PRT; 605 AA.
AC P87053;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof1 (Skp1-binding protein 1).
GN POF1 OR SBI1 OR SPAC57A10.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Katayama S., Kitamura K., Toda T.;
RT "Systematic genome-wide analysis of F-box protein-encoding genes in
RT fission yeast."

```

Submitted (SEP-1999) to the EMBL/GenBank/DBU databases.

Sequence from N.A.

Strain-972:

W004V, Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Brooks K., Peat N., Hayles J., Baker S., Basham D., Bowman S., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagals K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., O'Neill C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Stimmings M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkhardt G., Aert R., Robben J., Grymoprez B., Wellens I., Vanstreets E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Rappelli L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.

The genome sequence of *Schizosaccharomyces pombe*.

Nature 415:871-880(2002).

FUNCTION: Probably recognizes and binds to some phosphorylated proteins and promotes their ubiquitination and degradation.

PROTEIN: PART OF A SCF (SKP1-CUL1N-F-BOX) PROTEIN LIGASE COMPLEX (BY SIMILARITY).

1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.

1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL: AB032410; BAA84528.1; -

EMBL: Z94864; CAB08168.1; -

InterPro: IPR001810; F-box.

InterPro: IPR001680; WD40.

Pfam: PF00400; WD40; 7.

Pfam: PF00646; F-box; 1.

PRINTS: PR00320; GPOTENINRPT.

PRODOM: PD000018; WD40; 5.

SMART: SM00256; FBOX; 1.

SMART: SM00320; WD40; 7.

PROSITE: PS50181; FBOX; 1.

PROSITE: PS50678; WD\_REPEATS\_1; 2.

PROSITE: PS50082; WD\_REPEATS\_2; 7.

PROSITE: PS50294; WD\_REPEATS\_REGION; 1.

Ubl conjugation: Repeat; WD repeat.

DOMAIN 107 153

REPEAT 271 299 WD 1.

REPEAT 311 339 WD 2.

REPEAT 350 379 WD 3.

REPEAT 390 420 WD 4.

REPEAT 432 460 WD 5.

REPEAT 472 500 WD 6.

REPEAT 510 538 WD 7.

SEQUENCE 605 AA; 67110 MW; 711809379ECSCLF0 CRC64;

Query Match 45.4%; Score 69; DB 1; Length 605;

Best Local Similarity 51.9%; Pred. No. 0.0089;

Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

3 HSGRVERLQDFEFOIVSSSHDDITLW 29

DB 312 HSGGVICLQDFQCKLISGMDTIRW 338

RESULT 7

SC02\_NEUCR STANDARD; PRT; 650 AA.

AC 001277:

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Sulfur controller-2 (SCON2).

GN SCON-2.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales; Sordariaceae; Neurospora.

OX NCBI\_Taxid=5141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=74-OR23-1A; PubMed=7724564;

RX MEDLINE=95241499; PubMed=7724564;

RT Kumar A., Paletta J.V.

RT "The sulfur controller-2 negative regulatory gene of *Neurospora crassa* encodes a protein with beta-transducin repeats."

RT Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).

CC 1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.

CC 1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.

CC 1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).

CC 1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.

CC 1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL: U17251; AAA6968.1; -

InterPro: IPR001810; F-box.

InterPro: IPR001680; WD40.

Pfam: PF00400; WD40; 7.

Pfam: PF00646; F-box; 1.

PRINTS: PR00320; GPOTENINRPT.

PRODOM: PD000018; WD40; 4.

SMART: SM00256; FBOX; 1.

SMART: SM00320; WD40; 7.

PROSITE: PS50181; FBOX; 1.

PROSITE: PS50678; WD\_REPEATS\_1; 2.

PROSITE: PS50082; WD\_REPEATS\_2; 6.

PROSITE: PS50294; WD\_REPEATS\_REGION; 1.

KW Transcription regulation; Repeat; WD repeat.

DOMAIN 124 170

REPEAT 292 320 WD 1.

REPEAT 332 360 WD 2.

REPEAT 372 400 WD 3.

REPEAT 411 441 WD 4.

REPEAT 453 488 WD 5.

REPEAT 528 564 WD 6.

REPEAT 576 604 WD 7.

REPEAT 616 644 WD 8.

SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 43.4%; Score 66; DB 1; Length 650;

Best Local Similarity 41.4%; Pred. No. 0.028;

Matches 12; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

1 VHSGRVERLQDFEFOIVSSSHDDITLW 29







FT REPEAT 145 185 WD 2.  
 FT REPEAT 188 227 WD 3.  
 FT REPEAT 237 276 WD 4.  
 FT REPEAT 322 361 WD 5.  
 FT REPEAT 446 485 WD 6.  
 FT REPEAT 489 528 WD 7.  
 FT REPEAT 534 573 WD 8.  
 FT REPEAT 579 610 WD 9.  
 SO SEQUENCE 611 AA: 65323 MW: 821452C61B5D27A CRC64;

Query Match 40.8%; Score 62; DB 1; Length 611;  
 Best Local Similarity 51.6%; Pred. No. 0.1;  
 Matches 16; Conservative 3; Mismatches 10; Indels 2; Gaps 1;

Oy 1 VEHSGRVFRLQF--DEFOIVSSSHDDITLW 29  
 Db 236 VAHSGVFGLTWSPDGTKIASADKTIKIM 266

RESULT 12  
 Y1X1 SCHPO STANDARD; PRT; 651 AA.  
 AC Q9P713;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical WD-repeat protein C25D11.01 in chromosome I.  
 GN SPAC25D11.01  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 ON NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.-A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkert G., Aert R., Robben J., Grynopretz B.,  
 RA Weltjens I., Vansteede E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eber P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Gerton A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez J., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AL157993; CAB76232.1;  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 6.  
 DR PRINTS: PRO0320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS00682; WD\_REPEATS\_2; 5.  
 DR PROSITE: PS00294; WD\_REPEATS\_REGION; 1.  
 KM Hypothetical protein Repeat, WD repeat.  
 FT REPEAT 350 369 WD 1.  
 FT REPEAT 372 409 WD 2.  
 FT REPEAT 433 473 WD 3.  
 FT REPEAT 492 531 WD 4.  
 FT REPEAT 534 573 WD 5.  
 FT REPEAT 575 610 WD 6.  
 FT REPEAT 620 650 WD 7.  
 SQ SEQUENCE 651 AA: 72372 MW: 4F59D1104A50DBA CRC64;

Query Match 40.1%; Score 61; DB 1; Length 651;  
 Best Local Similarity 34.5%; Pred. No. 0.16;  
 Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Oy 1 VEHSGRVFRLQF--DEFOIVSSSHDDITLW 29  
 Db 533 IGH7APISLQFDSNHLGSGYDNSVRLW 561

RESULT 13  
 SCOB\_EMENT STANDARD; PRT; 678 AA.  
 ID SCOB\_EMENT  
 AC Q00659;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sulfur metabolite repression control protein.  
 GN SCONB OR MAPBI.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocomaceae; Emericella.  
 ON NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Natoff R.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE  
 CC REPRESSION.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT  
 CC PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Transcription regulation; Repeat; WD repeat.  
 KM DOMAIN 178 224 F-BOX.  
 FT REPEAT 178 224 F-BOX.  
 FT REPEAT 375 375 WD 1.  
 FT REPEAT 387 415 WD 2.  
 FT REPEAT 427 455 WD 3.  
 FT REPEAT 466 496 WD 4.  
 FT REPEAT 508 543 WD 5.  
 FT REPEAT 553 595 WD 6.  
 FT REPEAT 607 635 WD 7.  
 FT REPEAT 647 675 WD 8.  
 SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;  
 Query Match 40.1%; Score 61; DB 1; Length 678;  
 Best Local Similarity 40.7%; Pred. No. 0.17;  
 Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;  
 QY 3 HSGRVFLQFDFQVSSSHDTILIM 29  
 Db 388 HESGIRCLQFDQTKLISGMDRTIKW 414  
 RESULT 14  
 GBB3\_TOBAC STANDARD; PRT; 375 AA.  
 AC Q40507;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Guanine nucleotide-binding protein beta subunit.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Samsun NN; TISSUE=leaf;  
 RA Kustretsov V.V., Oelmueller R.;  
 RT "Isolation and characterization of cDNAs encoding the subunit beta of  
 heterotrimeric G proteins from G. tabacum.";  
 RT (In) Plant Gene Register PGR96-048.  
 CC  
 CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE  
 SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE  
 GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-  
 RECEPTOR INTERACTION.  
 CC  
 CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: X98161; CAA66842.1; -.  
 DR HSSP: P04901; ITRG.  
 DR InterPro: IPR001632; Gprotein\_B.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00319; GPROTEINB.  
 DR PRINTS: PR00320; GPROTEINBPT.  
 DR PRODOM: PD000018; WD40; 5.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 5.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Transducer; Repeat; WD repeat; Multigene family.  
 FT REPEAT 63 93 WD 1.  
 FT REPEAT 105 135 WD 2.  
 FT REPEAT 154 185 WD 3.

FT REPEAT 202 233 WD 4.  
 FT REPEAT 246 276 WD 5.  
 FT REPEAT 293 323 WD 6.  
 FT REPEAT 339 369 WD 7.  
 SQ SEQUENCE 375 AA; 40812 MW; 72D53D1DFC5CF598 CRC64;  
 Query Match 37.5%; Score 57; DB 1; Length 375;  
 Best Local Similarity 34.5%; Pred. No. 0.34;  
 Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 1;  
 QY 3 HSGRVFLQF--DFQVSSSHDTILIM 29  
 Db 64 HTGKVSIDWTPKRNIVSASODGRLIW 92  
 RESULT 15  
 GBB1\_TOBAC STANDARD; PRT; 377 AA.  
 AC P93397;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Guanine nucleotide-binding protein beta subunit 1.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. SRL; TISSUE=Leaf;  
 RA Lein W., Saalbach G.;  
 RT Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE  
 SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE  
 GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-  
 RECEPTOR INTERACTION.  
 CC  
 CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: Z84820; CAB06618.1; -.  
 DR HSSP: P04901; ITRG.  
 DR InterPro: IPR001632; Gprotein\_B.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00319; GPROTEINB.  
 DR PRINTS: PR00320; GPROTEINBPT.  
 DR PRODOM: PD000018; WD40; 5.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 5.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Transducer; Repeat; WD repeat; Multigene family.  
 FT REPEAT 63 93 WD 1.  
 FT REPEAT 105 135 WD 2.  
 FT REPEAT 154 185 WD 3.  
 FT REPEAT 202 233 WD 4.  
 FT REPEAT 246 276 WD 5.  
 FT REPEAT 293 323 WD 6.  
 FT REPEAT 339 369 WD 7.  
 SQ SEQUENCE 377 AA; 41145 MW; 23351986467A079B CRC64;  
 Query Match 37.5%; Score 57; DB 1; Length 377;  
 Best Local Similarity 34.5%; Pred. No. 0.34;

Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 1;

OY 3 HSGRVERLQF-DEFOIVSSSHDDTILIW 29

Db 64 HTGKVYSLDWTPEKNRIVSASQDGRLLIW 92

Search completed: April 10, 2003, 13:19:38  
Job time : 2.08404 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 11:23:10 : Search time 6.85262 Seconds  
(without alignments)  
871.983 Million cell updates/sec

Title: US-09-601-168b-2\_COPY\_516\_544

Perfect score: 152  
Sequence: 1 VEHSGRVRLQDFEFQIVSSSHDTILW 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBL.21:  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	100.0	569	11 09R1G7	09R1G7 mus musculus
2	152	100.0	569	11 09R159	09R159 mus musculus
3	152	100.0	569	11 09Q015	09Q015 mus musculus
4	151	99.3	252	11 09Q2C7	09Q2C7 mus musculus
5	151	99.3	563	11 09Q23H0	09Q23H0 mus musculus
6	149	98.0	510	5 044382	044382 drosophila
7	149	98.0	510	5 09YDE3	09YDE3 drosophila
8	141	92.8	430	5 09B154	09B154 heterodera
9	116	76.3	47	13 08UUN2	08UUN2 xenopus lae
10	73	48.0	553	4 09MDX6	09MDX6 homo sapien
11	73	48.0	561	4 096R12	096R12 homo sapien
12	73	48.0	589	4 096LE0	096LE0 homo sapien
13	73	48.0	627	4 096A16	096A16 homo sapien
14	73	48.0	629	11 08VHP4	08VHP4 mus musculus
15	73	48.0	629	11 08BVA4	08BVA4 mus musculus
16	73	48.0	707	4 0969H0	0969H0 homo sapien

17	73	48.0	1326	5 09VZP4	09VZP4 drosophila
18	69	45.4	293	4 08TC14	08TC14 homo sapien
19	69	45.4	468	11 09DB63	09DB63 mus musculus
20	69	45.4	472	4 096D26	096D26 homo sapien
21	69	45.4	472	4 09BR11	09BR11 homo sapien
22	69	45.4	504	4 096JY5	096JY5 homo sapien
23	65	42.8	391	5 09VU65	09VU65 drosophila
24	64.5	42.4	525	4 09BWS5	09BWS5 homo sapien
25	62	40.8	992	16 098GJ0	098GJ0 rhizobium
26	61	40.1	220	5 09VAG1	09VAG1 drosophila
27	61	40.1	419	5 09STJ1	09STJ1 enccephalito
28	61	40.1	556	5 08SC04	08SC04 enccephalito
29	59	38.8	727	5 08SS07	08SS07 xenopus lae
30	57	37.5	122	13 08UUN3	08UUN3 secenaromyc
31	57	37.5	330	3 005583	08Y22 phytophthor
32	57	37.5	344	10 08RYB2	08RYB2 nicotiana t
33	57	37.5	377	10 09FV61	09FV61 plasm sativ
34	57	37.5	377	10 09XFK0	09XFK0 plasm sativ
35	57	37.5	377	10 09SW94	09SW94 solanum tub
36	57	37.5	380	10 0945H7	0945H7 solanum tub
37	57	37.5	380	10 064944	064944 avena fatua
38	57	37.5	423	5 08SSJ1	08SSJ1 enccephalito
39	57	37.5	1221	4 09Y2K9	09Y2K9 homo sapien
40	56	36.8	311	10 09AVW0	09AVW0 guillardia
41	56	36.8	316	3 09HGV7	09HGV7 emericella
42	56	36.8	375	10 094AR3	094AR3 arabidopsis
43	56	36.8	496	5 093339	093339 caenorhabdi
44	56	36.8	505	4 08WUD8	08WUD8 homo sapien
45	56	36.8	513	11 055039	055039 mus musculus

## ALIGNMENTS

RESULT 1  
ID 09R1G7 PRELIMINARY: PRT; 569 AA.

AC 09R1G7; 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE beta-transducin repeat-containing protein.  
GN BTRC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Winston J., Elledge S.J., Harper J.W.;  
Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
CC -i- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
DR EMBL: AF110396; AADA1025.1; .  
DR MGD: MGI:1338871; BTRC.  
DR InterPro: IPR001810; F-box.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00646; F-box; 1.  
DR Pfam: PF00400; WD40; 7.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR PRODOM: PD000018; WD40; 4.  
DR SMART: SM00256; FBOX; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS00181; FBOX; 1.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
DR PROSITE: PS00682; WD\_REPEATS\_2; 7.  
DR PROSITE: PS00294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

Query Match 100.0%; Score 152; DB 11; Length 569;  
Best Local Similarity 100.0%; Pred. No. 7.8e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

GN FBXW1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: BC008552; AAC08552.1;
DR MGI: 2144023; Fbxw1b.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 6.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 3.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KM Repeat: WD repeat.
FT NON_TER
SQ SEQUENCE 252 AA; 28424 MW; F71737CBD7A9F75F CRC64;

Query Match 99.3%; Score 151; DB 11; Length 252;
Best Local Similarity 96.6%; Pred. No. 4.6e-15;
Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVRLQDFEFQIVSSSHDITLIW 29
Db 199 VEHSGRVRLQDFEFQIVSSSHDITLIW 227
|||||
ID 0923H0 PRELIMINARY; PRT; 563 AA.
AC 0923H0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F-box/WD40 repeat-containing protein HOS.
GN FBXW1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=EV/N;
RC Bhatia N., Hertler J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;
RT "Mouse homolog of HOS (MHOS) is overexpressed in skin tumors and
RT implicated in constitutive activation of NF-kappaB."
CC Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AY038079; AAK72095.1;
DR MGI: 2144023; Fbxw1b.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRODOM: PD000018; WD40; 4.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KM Repeat: WD repeat.
SQ SEQUENCE 563 AA; 64741 MW; 9A8562F3FF5E3496 CRC64;

Query Match 99.3%; Score 151; DB 11; Length 563;
Best Local Similarity 96.6%; Pred. No. 1.1e-14;
Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVRLQDFEFQIVSSSHDITLIW 29
|||||

```

```

Db 510 VEHSGRVRLQDFEFQIVSSSHDITLIW 538

RESULT 6
QY 044382 PRELIMINARY; PRT; 510 AA.
AC 044382;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE SLMB.
GN SLMB OR SLMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCB1_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98121115; PubMed=9461217;
RA Jiang J., Struhl G.;
RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-
RT box/WD40-repeat protein Slmb."
RL Nature 391:493-496(1998).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF032878; AAC38852.1;
DR FLYBase: FBgn0023423; slmb.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KM Repeat: WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; DBB0243D3730A5E8 CRC64;

Query Match 98.0%; Score 149; DB 5; Length 510;
Best Local Similarity 96.6%; Pred. No. 2e-14;
Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEHSGRVRLQDFEFQIVSSSHDITLIW 29
Db 459 VEHSGRVRLQDFEFQIVSSSHDITLIW 487
|||||

RESULT 7
QY 09VDE3 PRELIMINARY; PRT; 510 AA.
AC 09VDE3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE SLMB protein (SLMB).
GN SLMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCB1_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandides P.G., Scherf S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

```

Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 April J.F., Aphayani A., An H.-J., Andrews-Franko C., Baldwin D.,  
 Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borikova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,  
 Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 Glorides A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,  
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkllov G., Milshina N.V., Mobarry C., Morris J., Moshireli A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spralling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weltschbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Zhang R.A., Myers E.W., Rubin G.W., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-OVARY, AND IMAGINAL DISC;  
 RX MEDLINE=202455299; PubMed=10781936;  
 RA Miletich I., Limbourg-Bouchon B.;  
 RT "Drosophila null limb clones transiently deregulate Hedgehog-  
 independent transcription of wingless in all limb discs, and induce  
 RT decapentaplegic transcription linked to imaginal disc regeneration.";  
 RL Mech. Dev. 93:15-26(2000).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF003733; AAF5853.1; -;  
 DR EMBL: AF222924; AAF63214.1; -;  
 DR EMBL: AF222923; AAF63213.1; -;  
 DR FLYBase; FBgn0023423; limb.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00646; F-box; 1.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS50181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 510 AA: 58952 MW; FAD5DF126F58A012 CRC64;

Query Match 98.0%; Score 149; DB 5; Length 510;  
 Best Local Similarity 96.6%; Pred. No. 2e-14;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VHSGRVRLQDFEFOIVSSSHDITLIM 29  
 DB 459 VHTGRVRLQDFEFOIVSSSHDITLIM 487

RESULT 8  
 09B554

ID 09B554 PRELIMINARY; PRT; 430 AA.  
 AC 09B554;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE beta-transducin repeat-containing protein (Fragment).  
 OS Heterodera glycines (Soybean cyst nematode).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;  
 OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.  
 OX NCBI\_TaxID=51029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kovaleva E.S., Yakovlev A.G., Masler E.P.;  
 RT "Plant parasitic nematode b-TRCP.";  
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AF339101; AAK26376.1; -;  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR ProDom; PD000018; WD40; 4.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN.6.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER 1  
 SQ SEQUENCE 430 AA: 47916 MW; 4ABC3F2DFE3A50B CRC64;

Query Match 92.8%; Score 141; DB 5; Length 430;  
 Best Local Similarity 89.3%; Pred. No. 2.8e-13;  
 Matches 25; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 EHSGRVRLQDFEFOIVSSSHDITLIM 29  
 DB 341 QHTGRVRLQDFEFOIVSSSHDITLIM 368

RESULT 9  
 ID 080UN2 PRELIMINARY; PRT; 47 AA.  
 AC 080UN2;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Beta-TRCP protein (Fragment).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Carnevali F.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ballarino M.;  
 RT "Analisi strutturale e funzionale del gene beta-TRCP in *Xenopus*  
 RT laevis".  
 RT Thesis (2001).  
 RL Department of Genetics and Molecular Biology "Charles Darwin",  
 RL University of Rome La Sapienza, Rome, Italy.  
 DR EMBL; AJ428939; CAD21936.1; -;  
 DR InterPro; IPR001680; WD40.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN.1.  
 FT NON\_TER 1  
 SQ SEQUENCE 47 AA: 5521 MW; FD0A221906DC70B9 CRC64;

Query Match 76.3%; Score 116; DB 13; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```

AC 096A16;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Archipelago beta form (F-box protein FBW7).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21449047; PubMed=11565033;
RA Moberg K.H., Bell D.W., Maher D.C., Haber D.A., Hariharan I.K.;
RT "Archipelago regulates Cyclin E levels in Drosophila and is mutated in
RT human cancer cell lines."
RL Nature 413:311-316(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=20003061; PubMed=10531037;
RA Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.;
RT "A family of mammalian F-box proteins."
RL Curr. Biol. 9:1180-1182(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Koepf D.M., Winston J.T., Harper W., Ellledge S.J.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF411972; AAL06291.1; -
DR EMBL: AY033553; AKS7547.1; -
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR ProDom: PD000018; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ
SEQUENCE 627 AA; 70324 MW; 3D4107C053381BED CRC64;

Query Match
Best Local Similarity 48.0%; Score 73; DB 4; Length 627;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVFRLQDFEFOIVSSSHDITLIM 29
DB 460 HTNRVYSLOFDGIHVYSGSLDTSIRW 486

RESULT 14
QYVH4 PRELIMINARY; PRT; 629 AA.
AC QYVH4;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE F-box protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ilyin G.P.;
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF427101; AAL50052.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.

```

```

DR ProDom: PD000018; WD40; 2.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
SQ
SEQUENCE 629 AA; 70562 MW; BE916405A3490A3E CRC64;

Query Match
Best Local Similarity 48.0%; Score 73; DB 11; Length 629;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVFRLQDFEFOIVSSSHDITLIM 29
DB 462 HTNRVYSLOFDGIHVYSGSLDTSIRW 488

RESULT 15
QYVH4 PRELIMINARY; PRT; 629 AA.
AC QYVH4;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE F-box-WD40 repeat protein 6.
GN FBXW6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=129/SV;
RC MEDLINE=21601157; PubMed=11735228;
RA Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,
RA Nakayama K.-I.;
RT "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of
RT Caenorhabditis elegans SEL-10."
RL Genomics 78:214-222(2001).
DR EMBL: AF391202; AAL40930.1; JOINED.
DR EMBL: AF391193; AAL40930.1; JOINED.
DR EMBL: AF391194; AAL40930.1; JOINED.
DR EMBL: AF391195; AAL40930.1; JOINED.
DR EMBL: AF391196; AAL40930.1; JOINED.
DR EMBL: AF391197; AAL40930.1; JOINED.
DR EMBL: AF391198; AAL40930.1; JOINED.
DR EMBL: AF391199; AAL40930.1; JOINED.
DR EMBL: AF391200; AAL40930.1; JOINED.
DR EMBL: AF391201; AAL40930.1; JOINED.
DR EMBL: AF391202; AAL40928.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 2.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 8.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
SQ
SEQUENCE 629 AA; 70547 MW; 80483A7C2ED61FE7 CRC64;

Query Match
Best Local Similarity 48.0%; Score 73; DB 11; Length 629;
Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HSGRVFRLQDFEFOIVSSSHDITLIM 29
DB 462 HTNRVYSLOFDGIHVYSGSLDTSIRW 488

```

Fri Apr 11, 13:31:14 2003

us-09-601-168b-2\_copy\_516\_544.rspt

Page 7

Search completed: April 11, 2003, 11:51:55  
Job time : 7.85262 secs

---

